



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 113875

TO: Minh-Tam Davis  
Location: Rem 3a24 / 3c18  
Tuesday, February 10, 2004  
Art Unit: 1642  
Phone: 272-0830  
Serial Number: 10 / 017066

From: Jan Delaval  
Location: Biotech-Chem Library  
Rem 1A51  
Phone: 272-2504  
[jan.delaval@uspto.gov](mailto:jan.delaval@uspto.gov)

### Search Notes

priority date 10/2000  
10/1999

extra search requested 02/18/04

From: Chan, Christina  
Sent: Monday, February 09, 2004 2:09 PM  
To: Davis, Minh-Tam; STIC-Biotech/ChemLib  
Subject: RE: Rush search request for 10/017066

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644  
(571)-272-0841  
Remsen, 3E89

-----Original Message-----

From: Davis, Minh-Tam  
Sent: Monday, February 09, 2004 12:11 PM  
To: Chan, Christina  
Subject: Rush search request for 10/017066

Please search in commercial database, issued patent files, PGPUB and for interference:  
SEQ ID NO:1 and 2.

Thank you.

MINH TAM DAVIS  
ART UNIT 1642  
RESEM, ROOM 3A24, MB 3C18  
272-0830

Searcher:                       
Phone: 22504  
Location:                       
Date Picked Up: 215  
Date Completed: 216  
Searcher Prep/Review:                       
Flerical: 10  
Online time: 15

TYPE OF SEARCH:

NA Sequences:                       
AA Sequences:                       
Structures:                       
Bibliographic:                       
Litigation:                       
Full text:                       
Patent Family:                       
Other:                     

RECEIVED  
FEB 19 2004  
STIC/CHEN Division

VENDOR/COST (where applic.)  
STN:                       
DIALOG:                       
Questel/Orbit:                       
DRLink:                       
Lexis/Nexis:                       
Sequence Sys.:                       
WWW/Internet:                       
Other (specify):







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 19:20:07 ; Search time 1052 Seconds

(without alignment(s))  
10980.867 Million cell updates/sec

Title: US-10-017-066A-1

Perfect score: 3136  
Sequence: 1 ccgagagagcgtcgtatcctcagt.....aaaaaaaaaaaaaaaaa 3136

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

```
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3136	100.0	3136	15	US-10-001-469-2879 Sequence 2879, App
2	3136	100.0	3136	15	US-10-205-823-442 Sequence 442, App
3	3136	100.0	3136	15	US-10-225-567A-433 Sequence 433, App
4	2913.4	92.9	3104	12	US-10-295-027-872 Sequence 872, App
5	2028.4	64.7	2061	9	US-09-759-143-917 Sequence 917, App
6	2028.4	64.7	2061	9	US-09-780-669-917 Sequence 917, App
7	2028.4	64.7	2061	9	US-09-822-827-917 Sequence 917, App
8	2028.4	64.7	2061	10	US-09-895-793-917 Sequence 917, App
9	2028.4	64.7	2061	10	US-09-895-814-917 Sequence 917, App
10	2028.4	64.7	2061	13	US-10-144-678A-917 Sequence 917, App
11	2028.4	64.7	2061	13	US-10-294-025-917 Sequence 917, App
12	2028.4	64.7	2061	14	US-10-012-896-917 Sequence 917, App
13	1992.6	63.5	2034	11	US-09-966-459A-4 Sequence 4, Appl1
14	1287.8	41.1	1302	9	US-09-759-143-916 Sequence 916, App
15	1287.8	41.1	1302	9	US-09-780-669-916 Sequence 916, App

16	1287.8	41.1	1302	9	US-09-822-827-916 Sequence 916, App
17	1287.8	41.1	1302	10	US-09-895-793-916 Sequence 916, App
18	1287.8	41.1	1302	10	US-09-895-814-916 Sequence 916, App
19	1287.8	41.1	1302	13	US-10-144-678A-916 Sequence 916, App
20	1287.8	41.1	1302	13	US-10-294-025-916 Sequence 916, App
21	1287.8	41.1	1302	14	US-10-012-896-916 Sequence 916, App
22	1186.2	37.8	1354	12	US-10-292-798-745 Sequence 745, App
23	1186.2	37.8	1357	13	US-10-017-161-867 Sequence 867, App
24	966.4	30.8	968	13	US-10-044-643-25 Sequence 25, Appl1
25	964.2	30.7	969	13	US-10-044-643-27 Sequence 27, Appl1
26	963.2	30.7	968	13	US-10-044-643-29 Sequence 29, Appl1
27	963.4	30.5	957	9	US-09-759-143-918 Sequence 918, App
28	963.4	30.5	957	9	US-09-780-669-918 Sequence 918, App
29	963.4	30.5	957	9	US-09-822-827-918 Sequence 918, App
30	963.4	30.5	957	10	US-09-886-055-92 Sequence 92, Appl1
31	963.4	30.5	957	10	US-09-895-793-918 Sequence 918, App
32	963.4	30.5	957	10	US-09-895-814-918 Sequence 918, App
33	963.4	30.5	957	11	US-09-866-459A-1 Sequence 1, Appl1
34	963.4	30.5	957	11	US-09-804-291-92 Sequence 92, Appl1
35	963.4	30.5	957	13	US-10-144-678A-918 Sequence 918, App
36	963.4	30.5	957	13	US-10-294-025-918 Sequence 918, App
37	963.4	30.5	957	14	US-10-012-896-918 Sequence 918, App
38	963.4	30.4	954	9	US-09-759-143-919 Sequence 919, App
39	963.4	30.4	954	9	US-09-780-669-919 Sequence 919, App
40	963.4	30.4	954	9	US-09-822-827-919 Sequence 919, App
41	963.4	30.4	954	10	US-09-895-793-919 Sequence 919, App
42	963.4	30.4	954	10	US-09-895-814-919 Sequence 919, App
43	963.4	30.4	954	13	US-10-144-678A-919 Sequence 919, App
44	963.4	30.4	954	13	US-10-294-025-919 Sequence 919, App
45	963.4	30.4	954	13	US-10-029-386-20673 Sequence 20673, A

## ALIGNMENTS

```
RESULT 1
US-10-001-469-2879
; Sequence 2879, Application US/10001469
; Publication No. US20030091562A1
;
GENERAL INFORMATION:
; APPLICANT: JAKOBOVITS, AVA
; APPLICANT: RATTANO, ARTHUR
; APPLICANT: AFRAN, DANIEL
; APPLICANT: SAFRAN, DOUGLAS
; APPLICANT: HUBERT, RENE
; APPLICANT: PARIS, MARY
; APPLICANT: CHALITA-BID, PIA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED
; FILE REFERENCE: 51158-20024.20
; CURRENT APPLICATION NUMBER: US/10/001,469
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/157,902
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/291,118
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 09/680,728
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 2888
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2879
; LENGTH: 3136
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 101P3A11 cDNA
; NAME/KEY: CDS
; LOCATION: (130)..(1083)
US-10-001-469-2879
Query Match 100.0%; Score 3136; DB 15; Length 3136;
```



QY	2161	TTCTGATTTGGGCGATGAGGTTTAGGGAGCCACCAGATTATATATGGGAGCTATGGAAATGGCAGG	2220
Db	2161	TTCTGATTTGGGCGATGAGGTTTAGGGAGCCACCAGATTATATATGGGAGCTATGGAAATGGCAGG	2220
QY	2221	TTCTGAGATTAACATTTGGCTTTTGAAGTGTGACTGTGACCTGGAAAGTAGGGGATCTTC	2280
Db	2221	TTCTGAGATTAACATTTGGCTTTTGAAGTGTGACTGTGACCTGGAAAGTAGGGGATCTTC	2280
QY	2281	AGGACCATGCTTATTTTGGGGGCTTTGTGCAGATATGGAAACAGGGAATTGTGAACCAAGAA	2340
Db	2281	AGGACCATGCTTATTTTGGGGGCTTTGTGTGAGATATGGAAACAGGGAATTGTGAACCAAGAA	2340
QY	2341	GCAATCTGACTTAGGCATGGGAATCAGGCAATTTTGCTTGTAGGGGCTATTACCAAGG	2400
Db	2341	GCAATCTGACTTAGGCATGGGAATCAGGCAATTTTGCTTGTAGGGGCTATTACCAAGG	2400
QY	2401	TTAATAGGTTTCATCTTCAACAGGATATGACACAGAGTGTAAACCAAGAACTCAATAC	2460
Db	2401	TTAATAGGTTTCATCTTCAACAGGATATGACACAGAGTGTAAACCAAGAACTCAATAC	2460
QY	2461	AAATACATAAACATGTGATCATATATATGTATAGTTTCATTTCTTTTCAATCCTCAGG	2520
Db	2461	AAATACATAAACATGTGATCATATATATGTATAGTTTCATTTCTTTTCAATCCTCAGG	2520
QY	2521	TTCCCTGATATGAGATTTCCCTATTAACATGCTTTCATCCTCTTTTGTATGATATATATTT	2580
Db	2521	TTCCCTGATATGAGATTTCCCTATTAACATGCTTTCATCCTCTTTTGTATGATATATATTT	2580
QY	2581	GGAAATGCGCTATTATATCTTGATTTGCTGTGACGTGATGACCATAGAGGCACTGTT	2640
Db	2581	GGAAATGCGCTATTATATCTTGATTTGCTGTGACGTGATGACCATAGAGGCACTGTT	2640
QY	2641	TATTTATGAATGTCATCTCTGTTCATCATTTGACTGCTCTTTGTCTCATATTTGCCCC	2700
Db	2641	TATTTATGAATGTCATCTCTGTTCATCATTTGACTGCTCTTTGTCTCATATTTGCCCC	2700
QY	2701	AGCAAAAGTGCCTAGAACATATATAGTGTATGCTTGAACACCGGTTATTTTTCATCAAAAC	2760
Db	2701	AGCAAAAGTGCCTAGAACATATATAGTGTATGCTTGAACACCGGTTATTTTTCATCAAAAC	2760
QY	2761	TGATTCCTTCTGTCTCTGAAACATATGCGCAGCAATTTTTCAGCCTTCTTGAATGGGTA	2820
Db	2761	TGATTCCTTCTGTCTCTGAAACATATGCGCAGCAATTTTTCAGCCTTCTTGAATGGGTA	2820
QY	2821	TTATTAATTTCTGGGCTTAATCTTCCAAATGTAGTGAAGTGAACATATGCAATTTCTATAC	2880
Db	2821	TTATTAATTTCTGGGCTTAATCTTCCAAATGTAGTGAAGTGAACATATGCAATTTCTATAC	2880
QY	2881	CTGGCTCATTAACACCTCCCATGTGACGCTTTTCAATGTTGACATTAATATGTGACTTGGGA	2940
Db	2881	CTGGCTCATTAACACCTCCCATGTGACGCTTTTCAATGTTGACATTAATATGTGACTTGGGA	2940
QY	2941	AGCTATGTGTACACAGAGTAAATCACACAGAAAGCTGGATTTCTGAAGAAATCTGTGAGA	3000
Db	2941	AGCTATGTGTGTACACAGAGTAAATCACACAGAAAGCTGGATTTCTGAAGAAATCTGTGAGA	3000
QY	3001	GCCAAACCTCTGTCAATTTGCACTCCCACTTGTATTTGTATGACAGGCAAGTGTGATTAAGTA	3060
Db	3001	GCCAAACCTCTGTCAATTTGCACTCCCACTTGTATTTGTATGACAGGCAAGTGTGATTAAGTA	3060
QY	3061	AAATTAAGTACTATTTGTGTCAAGAAAAAATTTTGTGAAGAAATTTTGTGAAGAAATTTTGTGA	3120
Db	3061	AAATTAAGTACTATTTGTGTCAAGAAAAAATTTTGTGAAGAAATTTTGTGAAGAAATTTTGTGA	3120
QY	3121	AAAAAAAAAAAAAAAAA 3136	
Db	3121	AAAAAAAAAAAAAAAAA 3136	

```

? Publication No. US20030108963A1
? GENERAL INFORMATION:
? APPLICANT: Schlegel, Robert
? APPLICANT: Morahan, John E.
? APPLICANT: Endege, Wilson O.
? APPLICANT: Gannavarapu, Manjula
? APPLICANT: Gorbacheva, Bella
? APPLICANT: Hoerssch, Sebastian
? APPLICANT: Kamatkar, Shubhangi
? APPLICANT: Monsey, Angela M.
? APPLICANT: Glatic, Karen
? APPLICANT: Zhao, Xumei
? APPLICANT: Anderson, Dustin
? TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
? TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
? TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
? FILE REFERENCE: MXI-044
? CURRENT APPLICATION NUMBER: US/10/205,823
? CURRENT FILING DATE: 2002-07-25
? PRIOR APPLICATION NUMBER: 60/307,982
? PRIOR FILING DATE: 2001-07-25
? PRIOR APPLICATION NUMBER: 60/314,356
? PRIOR FILING DATE: 2001-08-22
? PRIOR APPLICATION NUMBER: 60/325,020
? PRIOR FILING DATE: 2001-09-25
? PRIOR APPLICATION NUMBER: 60/341,746
? PRIOR FILING DATE: 2001-12-12
? PRIOR APPLICATION NUMBER: 60/362,158
? PRIOR FILING DATE: 2002-03-05
? NUMBER OF SEQ ID NOS: 455
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 442
? LENGTH: 3136
? TYPE: DNA
? ORGANISM: Homo sapiens
US-10-205-823-442

Query Match      100.0%; Score 3136; DB 15; Length 3136;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAGAGCGTGATTTCAGTGCAGCCGCGACGACTCTTGTGAGAGAGACTGCAGCAAG 60
DB      1 CAGAGAGCGTGATTTCAGTGCAGCCGCGACGACTCTTGTGAGAGAGACTGCAGCAAG 60
QY      61 GGGGTCAACATTCCTTCACATACGAGTTGAGCCTCTACCTGCTGAGTGCACAGTTG 120
DB      61 GGGGTCAACATTCCTTCACATACGAGTTGAGCCTCTACCTGCTGAGTGCACAGTTG 120
QY      121 AGCTTCTTCAATGATGTGTGATCCCAATGGCAATGAAATCCAGTCTACATCTTCAATCTTA 180
DB      121 AGCTTCTTCAATGATGTGTGATCCCAATGGCAATGAAATCCAGTCTACATCTTCAATCTTA 180
QY      181 ATAAGGCTCCCTGGATTAGAGAAGGCTCAGTTCTGTGTGGCCTTCCCATGTGCTCCCTC 240
DB      181 ATAAGGCTCCCTGGATTAGAGAAGGCTCAGTTCTGTGTGGCCTTCCCATGTGCTCCCTC 240
QY      241 TACCTTATTCGCTGCTAGAGTAATTGACAATCATCTAATTTGTGCGGAGTGCACAGC 300
DB      241 TACCTTATTCGCTGCTAGAGTAATTGACAATCATCTAATTTGTGCGGAGTGCACAGC 300
QY      301 CTGATAGAGCCCATGTATATATTTCTTGATGCTTTCAGGACATGACATCCCATCTCC 360
DB      301 CTGATAGAGCCCATGTATATATTTCTTGATGCTTTCAGGACATGACATCCCATCTCC 360
QY      361 ACCCTATCCATGCCCCAAATGCTGGCCATCTTCTGTGTAATTCACACTACCATCCAGTTT 420
DB      361 ACCCTATCCATGCCCCAAATGCTGGCCATCTTCTGTGTAATTCACACTACCATCCAGTTT 420
QY      421 GATGCTTGTCTGTATACAGATTTTGGCATCCACTCCTTATCTGGATGGAATTCACAGTG 480
DB      421 GATGCTTGTCTGTATACAGATTTTGGCATCCACTCCTTATCTGGATGGAATTCACAGTG 480

```



Db	2641	AATATATGAANGTCATCTCTGTTCATCATTGACAGTCTCTTTGCTCATATTGAATATCCCCC	2700
Qy	2701	AGCAAAAGTGCTTAGAACCAATAATAGTCTTAATGCTTGACACCGGTATTTTTCATCAAAAC	2760
Db	2701	AGCAAAAGTGCTTAGAACCAATAATAGTCTTAATGCTTGACACCGGTATTTTTCATCAAAAC	2760
Qy	2761	TGATTTCCCTTCGTCTCGTAACAATAGCCAGGCAATTTTCCAGCCCTCTTTGAGTGGGTA	2820
Db	2761	TGATTTCCCTTCGTCTCGTAACAATAGCCAGGCAATTTTCCAGCCCTCTTTGAGTGGGTA	2820
Qy	2821	TTATTTAAATTTCTGGCCATTACTTCCATATGTGAGTGGAAAGTGACATGTGCATTTCTATAC	2880
Db	2821	TTATTTAAATTTCTGGCCATTACTTCCATATGTGAGTGGAAAGTGACATGTGCATTTCTATAC	2880
Qy	2881	CTGGCTCATAAACCCCTCCATATGTCAGACCTCTTTCATGTGTGACATTAATATGTGACTTGGGA	2940
Db	2881	CTGGCTCATAAACCCCTCCATATGTCAGACCTCTTTCATGTGTGACATTAATATGTGACTTGGGA	2940
Qy	2941	AGCATATGTTGTACACAGAGTAAATACCAAGAAAGCTGGATTTCTGAAAAAACTGTGCAGA	3000
Db	2941	AGCATATGTTGTACACAGAGTAAATACCAAGAAAGCTGGATTTCTGAAAAAACTGTGCAGA	3000
Qy	3001	GCCAAACCTCTGTGATTTGGCAACTCCCACTTGTAATTTGTACAGAGCAGTTGGAATAGTGA	3060
Db	3001	GCCAAACCTCTGTGATTTGGCAACTCCCACTTGTAATTTGTACAGAGCAGTTGGAATAGTGA	3060
Qy	3061	AAATTAAGTACTCTATGTGTGTCAGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA	3120
Db	3061	AAATTAAGTACTCTATGTGTGTCAGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA	3120
Qy	3121	AAAAAAAAAAAAAAAAAAAA 3136	
Db	3121	AAAAAAAAAAAAAAAAAAAA 3136	

RESULT 3  
US-10-225-567A-433  
Sequence 433, Application US/10225567A  
Publication No. US20030113798A1  
GENERAL INFORMATION:  
APPLICANT: Lifespan Biosciences  
APPLICANT: Brown, Joseph P.  
APPLICANT: Burner, Glenna C.  
APPLICANT: Roush, Christine L.  
TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR  
FILE REFERENCE: 1920-4-4  
CURRENT APPLICATION NUMBER: US/10/225,567A  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/257,144  
PRIOR FILING DATE: 2000-12-19  
NUMBER OF SEQ ID NOS: 2292  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 433  
LENGTH: 3136  
TYPES: DNA  
ORGANISM: Homo sapiens  
US-10-225-567A-433

	Query Match	Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
QY	1	CAGAGAGGCTTATTTTCAGTCAGAGCTCCAGACCTCTTCTGGAGAGAGCTGACCAAG	60							
Db	1	CAGAGAGGCTTATTTTCAGTCAGAGCTCCAGACCTCTTCTGGAGAGAGCTGACCAAG	60							
QY	61	GGGCTACACATTCCTTCATACGAGTTGAGCCTCTACCTGCTGGTGTGTCACAGTTC	120							
Db	61	GGGCTACACATTCCTTCATACGAGTTGAGCCTCTACCTGCTGGTGTGTCACAGTTC	120							
QY	121	AGCTCTTCATGATGGTGGATGCCAATGGCAATGAATCCAGTGTACATACCTTACCTCTTA	180							
Db	121	AGCTCTTCATGATGGTGGATGCCAATGGCAATGAATCCAGTGTACATACCTTACCTCTTA	180							

QY	181	ATPAGGCTCCCTGGTTTAAAGAGGCTCAGTTCGTGTGGCTTCGCATGTGGCTCCTC	240
Db	181	ATPAGGCTCCCTGGTTTAAAGAGGCTCAGTTCGTGTGGCTTCGCATGTGGCTCCTC	240
QY	241	TACCTTATGTGCTGTAGGTAACTTGACATCATCTPACATTGTGGGACTGACACAGC	300
Db	241	TACCTTATGTGCTGTAGGTAACTTGACATCATCTPACATTGTGGGACTGACACAGC	300
QY	301	CTGCATGAGCCCATGATATATTTCTTGTGATGTGTTCAAGCATTTGACATCTCATCTCC	360
Db	301	CTGCATGAGCCCATGATATATTTCTTGTGATGTGTTCAAGCATTTGACATCTCATCTCC	360
QY	361	ACCTCATTCATGCCCCAAATATGTGGCCATCTTCTGTTCATTCCACTACCATCAAGTTT	420
Db	361	ACCTCATTCATGCCCCAAATATGTGGCCATCTTCTGTTCATTCCACTACCATCAAGTTT	420
QY	421	GATGCTGTGTCTGATACAGATTTTGGCATCACTCCATTATCTGGCATGGAATCCACAGTG	480
Db	421	GATGCTGTGTCTGATACAGATTTTGGCATCACTCCATTATCTGGCATGGAATCCACAGTG	480
QY	481	CTGCTGGCCATGCTTTTGAACCGCTATGTGGCCATCTGTCAACCATGCGCCATGCCACA	540
Db	481	CTGCTGGCCATGCTTTTGAACCGCTATGTGGCCATCTGTCAACCATGCGCCATGCCACA	540
QY	541	GTACTTACGTTGCTCGTGTGCACCAAAATGTGTGTGCTGTGTGTGTGGGGGGGCTGCA	600
Db	541	GTACTTACGTTGCTCGTGTGCACCAAAATGTGTGTGCTGTGTGTGTGGGGGGGCTGCA	600
QY	601	CTGATGGACCCCTTCTGTCTTTCATCAAGCAGCTGCGCTTGTGCGCTCCCAATATCTT	660
Db	601	CTGATGGACCCCTTCTGTCTTTCATCAAGCAGCTGCGCTTGTGCGCTCCCAATATCTT	660
QY	661	TCCCATTCCTACGTGCTGCACCAAGATGATGAAGCTGGCGCTGTGATATCCGAGTC	720
Db	661	TCCCATTCCTACGTGCTGCACCAAGATGATGAAGCTGGCGCTGTGATATCCGAGTC	720
QY	721	AATGTGCTATATGGCCTTATGTCTATCATCATCTCGGCATTGAGCTGTGACTCATCTCATC	780
Db	721	AATGTGCTATATGGCCTTATGTCTATCATCATCTCGGCATTGAGCTGTGACTCATCTCATC	780
QY	781	TCCCTTCATATCGCTTATCTTAAACATGTGTGTGGAGCTTGACACGTGAAGCCACAGCC	840
Db	781	TCCCTTCATATCGCTTATCTTAAACATGTGTGTGGAGCTTGACACGTGAAGCCACAGCC	840
QY	841	AAGGCATTTGGCACTTGCTCTCATGTGTGTGCTGTGTTCATATCTATGACCTTTC	900
Db	841	AAGGCATTTGGCACTTGCTCTCATGTGTGTGCTGTGTTCATATCTATGACCTTTC	900
QY	901	ATTGATATTTCATATGTGATATCTGTTTGAAGAGGGGGTGAATCTCGCGCCCGCTATC	960
Db	901	ATTGATATTTCATATGTGATATCTGTTTGAAGAGGGGGTGAATCTCGCGCCCGCTATC	960
QY	961	TTGGCCAAATATCTATGCTGTGTCTCTCTGTGTCTACCCCAATGTCTATGTGATGAAG	1020
Db	961	TTGGCCAAATATCTATGCTGTGTCTCTCTGTGTCTACCCCAATGTCTATGTGATGAAG	1020
QY	1021	ACAAAGAGATATGACAGCGCATCTTTCGACTTTCATGTGGGCAACACAGCGCTTCAGAG	1080
Db	1021	ACAAAGAGATATGACAGCGCATCTTTCGACTTTCATGTGGGCAACACAGCGCTTCAGAG	1080
QY	1081	CCCTAGGTGCTAGTATCAAACTTCTTTCATATCAGAGTCTGTGATTCAGATTTTAAAT	1140
Db	1081	CCCTAGGTGCTAGTATCAAACTTCTTTCATATCAGAGTCTGTGATTCAGATTTTAAAT	1140
QY	1141	GTTAAACATTTTGGAAAGACAGTATCCAAAAAAATTTCTTAAATAAAAAATACATCTCA	1200
Db	1141	GTTAAACATTTTGGAAAGACAGTATCCAAAAAAATTTCTTAAATAAAAAATACATCTCA	1200
QY	1201	GATCCTTCAATATGAACCTGTGTGGGAATCTTCATTTTTCATATATATTTCTTCTT	1260
Db	1201	GATCCTTCAATATGAACCTGTGTGGGAATCTTCATTTTTCATATATATTTCTTCTT	1260



```

FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 872
LENGTH: 3104
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified base
LOCATION: (1)..(3104)
OTHER INFORMATION: n = g, a, c or t
US-10-295-027-872

Query Match      92.9%; Score 2913.4; DB 12; Length 3104;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2946; Conservative 1; Mismatches 18; Indels 3; Gaps 2;

```

```

481 CTGCTGGCCATGAGTTTGAACCGTATGTGGCCATCTGTACCCACCTGGCCATTCGCACA 540
541 GTACTTACGTTGGCTCTGTACCAAAATTTGTGTGCTGTGTGTGCGGGGGGCTGCA 600
541 GTACTTACGTTGGCTCTGTACCAAAATTTGTGTGCTGTGTGTGCGGGGGGCTGCA 600
601 CTGATGGACCCCTCTCTGTCTTCATCAAGACAGTGGCCCTTGCCGCTCAATATCTT 660
601 CTGATGGACCCCTCTCTGTCTTCATCAAGACAGTGGCCCTTGCCGCTCAATATCTT 660
661 TCCCATCTCTAAGCTTCAACCAAGATGTCAAGAGTGGCTGTGTATATTCGGGTC 720
661 TCCCATCTCTAAGCTTCAACCAAGATGTCAAGAGTGGCTGTGTATATTCGGGTC 720
721 AATGCTGTATGAGCTTATGTATCATATCGGCAATGGCCGAGCTCACTTCATC 780
721 AATGCTGTATGAGCTTATGTATCATATCGGCAATGGCCGAGCTCACTTCATC 780
781 TCCCTTCATATCTGCTTATTTTAAAGACTGTGGCTTGAACAGTGAAGCCAGGCC 840
781 TCCCTTCATATCTGCTTATTTTAAAGACTGTGGCTTGAACAGTGAAGCCAGGCC 840
841 AAGCATTTGGCACTTGGCTCTCATATGTGTGCTGTGTCTATTTATATGATACCTTC 900
841 AAGCATTTGGCACTTGGCTCTCATATGTGTGCTGTGTCTATTTATATGATACCTTC 900
901 ATTGATTTGTCATATGTGATCTGCTTTAGCAAGCGGAGTCTCTCGGCTCCGATC 960
901 ATTGATTTGTCATATGTGATCTGCTTTAGCAAGCGGAGTCTCTCGGCTCCGATC 960
961 TTGGCCATATCTATCTGTGTGTCTCTCTGTGTCTCAACCCATTTGTATGTAGAGAG 1020
961 TTGGCCATATCTATCTGTGTGTCTCTCTGTGTCTCAACCCATTTGTATGTAGAGAG 1020
1021 ACAAGAGAGATGTGACAGCGCATCTTGCACTTTCCATGTGGCCACACAGCTTGAG 1080
1021 ACAAGAGAGATGTGACAGCGCATCTTGCACTTTCCATGTGGCCACACAGCTTGAG 1080
1081 CCTTAGGTGTCATATGTATCAAACTTCTTTCATCAAGAGCTCTGATTAAGATTTTAT 1140
1081 CCTTAGGTGTCATATGTATCAAACTTCTTTCATCAAGAGCTCTGATTAAGATTTTAT 1140
1141 GTTAACTTTTGAAGAGAGATTTTCAAGAAAAAATTTCTTATATATAAATCACTCA 1200
1141 GTTAACTTTTGAAGAGAGATTTTCAAGAAAAAATTTCTTATATATAAATCACTCA 1200
1141 GTTAACTTTTGAAGAGAGATTTTCAAGAAAAAATTTCTTATATATAAATCACTCA 1200
1201 GATCTTCAAAATATGAATCTGTGGGGAATCTCAATTTTTCATATATTTTCTTCT 1260
1201 GATCTTCAAAATATGAATCTGTGGGGAATCTCAATTTTTCATATATTTTCTTCT 1260
1261 TGTCTTCTGTCATATATATATTAATTAATCCCTGATAGGTTGTGGAGGCTTATTA 1320
1261 TGTCTTCTGTCATATATATATTAATTAATCCCTGATAGGTTGTGGAGGCTTATTA 1320
1260 TGTCTTCTGTCATATATATTAATTAATCCCTGATAGGTTGTGGAGGCTTATTA 1319
1321 CTCTTCAATTTTACATGAGTCCAAATCTAAATCTGTTCTACTGATGTTTACAGATTC 1380
1321 CTCTTCAATTTTACATGAGTCCAAATCTAAATCTGTTCTACTGATGTTTACAGATTC 1380
1320 CTCTTCAATTTTACATGAGTCCAAATCTAAATCTGTTCTACTGATGTTTACAGATTC 1379
1381 TGAATTAAGAAATGTATCATCTAGAGAAATTTTGGCCAAAGCCCTTAACCAAGGAA 1440
1381 TGAATTAAGAAATGTATCATCTAGAGAAATTTTGGCCAAAGCCCTTAACCAAGGAA 1440
1380 TGAATTAAGAAATGTATCATCTAGAGAAATTTTGGCCAAAGCCCTTAACCAAGGAA 1439
1441 ATAAACAGAAATATATAAATATGATATATAGCTTAAATCTATATCTCTCTTCAG 1500
1441 ATAAACAGAAATATATAAATATGATATATAGCTTAAATCTATATCTCTCTTCAG 1499
1501 AACTCCCAACCAATGTGATCTCAAGAAAAATGTGTCTTCAAAATGACTTTCACAGAA 1560
1501 AACTCCCAACCAATGTGATCTCAAGAAAAATGTGTCTTCAAAATGACTTTCACAGAA 1559
1561 GAATATATTTTCTGTGACATAGCACTTAAGGGAAGATTTGAAGCTTTGAA 1620
1561 GAATATATTTTCTGTGACATAGCACTTAAGGGAAGATTTGAAGCTTTGAA 1620

```



Db 1560 GAATATATTTTCTCTGGACACTAGCACTTAAGGGGAAGATTGGAATTAAGCTTGAA 1619  
 QY 1621 AAGATTAATTTTACTAGCTTATGAAAAGTTGACACACTGTTCTGAGAATTTTCAACA 1680  
 Db 1620 AAGATTAATTTTACTAGCTTATGAAAAGTTGACACACTGTTCTGAGAATTTTCAACA 1679  
 QY 1681 TATGACCTGTTTCTTATTAATTTTCTTATTAACCTTTTAAATTAAGGAAAGATTT 1740  
 Db 1680 TATGACCTGTTTCTTATTAATTTTCTTATTAACCTTTTAAATTAAGGAAAGATTT 1739  
 QY 1741 ATTAATGACCTTATTAATTTTCTTATTAACCTTTTAAATTAAGGAAAGATTT 1800  
 Db 1740 ATTAATGACCTTATTAATTTTCTTATTAACCTTTTAAATTAAGGAAAGATTT 1799  
 QY 1801 TGGGCTCATCAAGATTAATAAATTAATAAATAAAGCTTCAAGCCCAATCTATGTA 1860  
 Db 1800 TGGGCTCATCAAGATTAATAAATTAATAAATAAAGCTTCAAGCCCAATCTATGTA 1857  
 QY 1861 TGTGGAAGAACTGTTAGAGAGCAACAAGGATGAGTTAGAGATTTCCAGAGCTTAC 1920  
 Db 1858 TGTGGAAGAACTGTTAGAGAGCAACAAGGATGAGTTAGAGATTTCCAGAGCTTAC 1917  
 QY 1921 ATTTCTAGAGAGATTTAATTTCTTCTCACTCATCAAGTGTGTAATTAAGAAATTC 1980  
 Db 1918 ATTTCTAGAGAGATTTAATTTCTTCTCACTCATCAAGTGTGTAATTAAGAAATTC 1977  
 QY 1981 CTGGCAAGAACTCATGAGCTTTAATCCCACTAGCTATGCTTAATGCTGCTGCAAT 2040  
 Db 1978 CTGGCAAGAACTCATGAGCTTTAATCCCACTAGCTATGCTTAATGCTGCTGCAAT 2037  
 QY 2041 GCGAATTAACCTGCTGCTTGAAGAAAGATTTCTAGGTTCAACATTAAGAAATTTCT 2100  
 Db 2038 GCGAATTAACCTGCTGCTTGAAGAAAGATTTCTAGGTTCAACATTAAGAAATTTCT 2097  
 QY 2101 TTCAGAAAGTCTGATAGGAGCTTATAGCAAGTATTAATTTTAAAGTCCAAAGTGA 2160  
 Db 2098 TTCAGAAAGTCTGATAGGAGCTTATAGCAAGTATTAATTTTAAAGTCCAAAGTGA 2157  
 QY 2161 TTTGATAGGAGAGTGAAGTGAAGGACCAAGTTATGATGGAAGTATGGAATGGCAG 2220  
 Db 2158 TTTGATAGGAGAGTGAAGTGAAGGACCAAGTTATGATGGAAGTATGGAATGGCAG 2217  
 QY 2221 TTTGATAGGAGAGTGAAGTGAAGGACCAAGTTATGATGGAAGTATGGAATGGCAG 2280  
 Db 2218 TTTGATAGGAGAGTGAAGTGAAGGACCAAGTTATGATGGAAGTATGGAATGGCAG 2277  
 QY 2281 AGGACCAAGTGAAGTGAAGGACCAAGTTATGATGGAAGTATGGAATGGCAG 2340  
 Db 2278 AGGACCAAGTGAAGTGAAGGACCAAGTTATGATGGAAGTATGGAATGGCAG 2337  
 QY 2341 GCAATCTGATAGGAGAGTGAAGGACCAAGTTATGATGGAAGTATGGAATGGCAG 2400  
 Db 2338 GCAATCTGATAGGAGAGTGAAGGACCAAGTTATGATGGAAGTATGGAATGGCAG 2397  
 QY 2401 TTAATAGTTCATCTTCAACAGATATGACACAGTGTATACCAAGAACTCAATTA 2460  
 Db 2398 TTAATAGTTCATCTTCAACAGATATGACACAGTGTATACCAAGAACTCAATTA 2457  
 QY 2461 AAATTAATCAAGTATCAATATGATGTAAGTTCATTTCTTTTCAATCTCAG 2520  
 Db 2458 AAATTAATCAAGTATCAATATGATGTAAGTTCATTTCTTTTCAATCTCAG 2517  
 QY 2521 TTTCCGTATGATGATCTTCAATAGATGTTTCAATCTTCTTGAATGATCAATTT 2580  
 Db 2518 TTTCCGTATGATGATCTTCAATAGATGTTTCAATCTTCTTGAATGATCAATTT 2577  
 QY 2581 GGAATGCTTATTAATCTTGAATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2640  
 Db 2578 GGAATGCTTATTAATCTTGAATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2637  
 QY 2641 TATTAATGATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
 Db 2638 TATTAATGATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2697

QY 2701 AGCAAGTCTTACACATATATAGTCTATGCTTACACACCGGTATTTTCAATCAAC 2760  
 Db 2698 AGCAAGTCTTACACATATATAGTCTATGCTTACACACCGGTATTTTCAATCAAC 2757  
 QY 2761 TGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
 Db 2758 TGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2817  
 QY 2821 TTAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
 Db 2818 TTAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2877  
 QY 2881 CTGGCTCATTAACCTCCCATGCTGAGCTTCAATGTTGATTAATGATGCTGGA 2940  
 Db 2878 CTGGCTCATTAACCTCCCATGCTGAGCTTCAATGTTGATTAATGATGCTGGA 2937  
 QY 2941 AGCTATGCTTACAGAGATTAATCAAC 2968  
 Db 2938 AGCTATGCTTACAGAGATTAATCAAC 2965

# RESULT 5 US-09-759-143-917

; Sequence 917, Application US/09759143  
 ; Patent No. US2002002248A1

## GENERAL INFORMATION:

; APPLICANT: Xu, Jianshun  
 ; APPLICANT: Dillon, David C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Ketter, Marc W.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Li, Samuel  
 ; APPLICANT: Mang, Ajun  
 ; APPLICANT: Skeiky, Yahir A.W.  
 ; APPLICANT: Hepler, William  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE REFERENCE: 210121.427C3  
 ; CURRENT APPLICATION NUMBER: US/09/759,143  
 ; CURRENT FILING DATE: 2001-01-12  
 ; NUMBER OF SEQ ID NOS: 934  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 917  
 ; LENGTH: 2061  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-759-143-917

## Query Match

Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1029 GATTGACAGGAGCACTTTCAGCTTTTCATGAGGAGCAACAGGCTTACAGGCTTAC 1088  
 Db 3 GATTGACAGGAGCACTTTCAGCTTTTCATGAGGAGCAACAGGCTTACAGGCTTAC 62  
 QY 1089 GTGAGTGAATCAACTTCTTTTCATTCAGAGTCTCTGATTCAGATTTTAATGTTAAT 1148  
 Db 63 GTGAGTGAATCAACTTCTTTTCATTCAGAGTCTCTGATTCAGATTTTAATGTTAAT 122  
 QY 1149 TTTGGAAGAGATTCAGAAATTAATTTCTTAAATTAATTAATTAATTAATTAATTA 1208  
 Db 123 TTTGGAAGAGATTCAGAAATTAATTTCTTAAATTAATTAATTAATTAATTAATTA 181



QY	1209	AAATATGAACCGTGTGGGAAATCCCAATTTTTCATATTAATTTCTCTTGTTCT	1268
Db	182	AAATATGAACCGTGTGGGAAATCCCAATTTTTCATATTAATTTCTCTTGTTCT	241
QY	1269	TGCTACATATATATTAATACCCTGACTAGGTTGTGGTAGGGATTAATTAATCTTTGCAT	1328
Db	242	TGCTACATATATATTAATACCCTGACTAGGTTGTGGTAGGGATTAATTAATCTTTGCAT	301
QY	1329	TTTACCATGACGTCCCAATCTTAACTGTCTTCACTGATGGTTTACAGCATTTCTGAGATTA	1388
Db	302	TTTACCATGACGTCCCAATCTTAACTGTCTTCACTGATGGTTTACAGCATTTCTGAGATTA	361
QY	1389	GAATGTACATCTAGAGAACATTTTGCCAAGGCGCTTAACGAGGCAAGGAAATTAACAC	1448
Db	362	GAATGTACATCTAGAGAACATTTTGCCAAGGCGCTTAACGAGGCAAGGAAATTAACAC	421
QY	1449	AGAAATATATATTAATGAGATTAATCTAGCTTAAACTATTAATCTTCTTCCAGACTCCCA	1508
Db	422	AGAAATATATATTAATGAGATTAATCTAGCTTAAACTATTAATCTTCTTCCAGACTCCCA	481
QY	1509	ACCACATTTGATCTCAGAAAAATGTGTCTCAAAATACCTTCAACAGAGAAATATAT	1568
Db	482	ACCACATTTGATCTCAGAAAAATGTGTCTTCAAAATACCTTCAACAGAGAAATATAT	541
QY	1569	TTTTCTCTGACACTAGCACTTAAGGGGAAGATTGAAAGTAAAGCCTTGAAAAAGATAC	1628
Db	542	TTTTCTCTGACACTAGCACTTAAGGGGAAGATTGAAAGTAAAGCCTTGAAAAAGATAC	601
QY	1629	ATTATACCTAGCTTATAGAAAGTGCACACATGTTCTGAGATTTTCAACAGATATGACCC	1688
Db	602	ATTATACCTAGCTTATAGAAAGTGCACACATGTTCTGAGATTTTCAACAGATATGACCC	661
QY	1689	CTGTATTTTCCATATTAATTTTCTTATCAACCCCTTAAATTAAGCAAGATATTAATAGAC	1748
Db	662	CTGTATTTTCCATATTAATTTTCTTATCAACCCCTTAAATTAAGCAAGATATTAATAGAC	721
QY	1749	CCTCATTTGACCATGGGAAATTCATGTTCACTGAGGGATCACTGAAATTAATGGGCTCA	1808
Db	722	CCTCATTTGACCATGGGAAATTCATGTTCACTGAGGGATCACTGAAATTAATGGGCTCA	781
QY	1809	TACAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1868
Db	782	TACAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	839
QY	1869	AACTGTTAGAGACCAACAGGGTAGTGGGTTAGAGATTTCCAGAGCTTTCATTTCTA	1928
Db	840	AACTGTTAGAGACCAACAGGGTAGTGGGTTAGAGATTTCCAGAGCTTTCATTTCTA	899
QY	1929	GAGGAGGATTTAATTTCTTCTCATCTCAATCCAGTGTGTAATTTAGGAATTTCCGCGAAC	1988
Db	900	GAGGAGGATTTAATTTCTTCTCATCTCAATCCAGTGTGTAATTTAGGAATTTCCGCGAAC	959
QY	1989	AGAACTCAGGCTTTAATCCCACTAGCTATATGCTTATGTCCTGTCGAATTTGCCAATTA	2048
Db	960	AGAACTCAGGCTTTAATCCCACTAGCTATATGCTTATGTCCTGTCGAATTTGCCAATTA	1019
QY	2049	CCCTGTCTTGGAGAAAGTGAATTTCTAGAGTCAACCATTAATGGAAGATTTCTTATCAGAAA	2108
Db	1020	CCCTGTGTCTTGGAGAAAGTGAATTTCTAGAGTCAACCATTAATGGAAGATTTCTTATCAGAAA	1079
QY	2109	GTCGTGCAATGAGGCTTATAGCAAGTAAATTTATTTTAAAGTTCCATAGGTATCTGATTA	2168
Db	1080	GTCGTGCAATGAGGCTTATAGCAAGTAAATTTATTTTAAAGTTCCATAGGTATCTGATTA	1139
QY	2169	GGCAGTGAAGTTAGGAGCCACCAAGTTATGATGGGAAGTATGAAATGGCAGGTCTTGAAG	2228
Db	1140	GGCAGTGAAGTTAGGAGCCACCAAGTTATGATGGGAAGTATGAAATGGCAGGTCTTGAAG	1199
QY	2229	ATAACATTTGGCCTTTTGAATGTGACTGTAATCTGTAAGGAAATCTTCAAGACAT	2288
Db	1200	ATAACATTTGGCCTTTTGAATGTGACTGTAATCTGTAAGGAAATCTTCAAGACAT	1259
QY	2289	GCTTTATTTGGGGCTTTGTGTGCAATATGGAACAGGGAATTTTGAACAGAAAGCAATCTG	2348

Db	1260	GCCTTAATTTG3GGCTTTGTGTGACATATGAAACAGGACACTTTAGACCAAGAAACGAAATCTG	1315
Qy	2349	ACTTAGGCATGGGAATCAGGACATTTTGGCTTCTGAGGGCTATTAACAAAGGTTTAATAG	2408
Db	1330	ACTTAGGCATGGGAATCAGGACATTTTGGCTTCTGAGGGGCTATTAACAAAGGTTTAATAG	1379
Qy	2409	TTTTCATCTTCAACAGGATATGACAAACAGTGTAAACAAAGAACTGTAATTCATAATCTA	2466
Db	1380	TTTTCATCTTCAACAGGATATGACAAACAGTGTAAACAAAGAACTGTAATTCATAATCTA	1433
Qy	2469	AAACATGTGATCATATATATGTGTGAAGTTTCATTTCTTTTCCAAATCCCTCAGAGTCCCTGA	2528
Db	1440	AAACATGTGATCATATATATGTGTGAAGTTTCATTTCTTTTCCAAATCCCTCAGAGTCCCTGA	1499
Qy	2529	TATGGATTCCTATAACATGCTTTTCATCCCTTTTGTAAATGGATATCATATTTTGGAAATGC	2588
Db	1500	TATGGATTCCTATAACATGCTTTTCATCCCTTTTGTAAATGGATATCATATTTTGGAAATGC	1555
Qy	2589	CTATTTAATACTGTATATTTGGTGTGTGACAGTGTAGCCCAATAGGGCACTGTTTATATAG	2644
Db	1560	CTATTTAATACTGTATATTTGGTGTGTGACAGTGTAGCCCAATAGGGCACTGTTTATATAG	1615
Qy	2649	AATGTCAATCTCTGTTCATCATTTGACGTCTCTTTGCTCATCATTTGAATCCCCAGCAAGT	2708
Db	1620	AATGTCAATCTCTGTTCATCATTTGACGTCTCTTTGCTCATCATTTGAATCCCCAGCAAGT	1679
Qy	2709	GCCTAGAACATATATAGTGTATATGCTTGTGACACCGGTTATTTTTCATCAAACTGTATCTT	2768
Db	1680	GCCTAGAACATATATAGTGTATATGCTTGTGACACCGGTTATTTTTCATCAAACTGTATCTT	1739
Qy	2769	TCTGTCTCTGAACACATATAGCAGGCAATTTTCCAGCCTTCTTGTAGTGGGATATATTA	2828
Db	1740	TCTGTCTCTGAACACATATAGCAGGCAATTTTCCAGCCTTCTTGTAGTGGGATATATTA	1799
Qy	2829	TTCTGGCCATTACTTCCATATGTGTAGTGTGAAGTGCATGTGTCAATTTCTATACCTGGCTCA	2888
Db	1800	TTCTGGCCATTACTTCCATATGTGTAGTGTGAAGTGCATGTGTCAATTTCTATACCTGGCTCA	1855
Qy	2889	TAAACCCCTCCCATATGTGTGACAGCCTTTCATATGTGAATTAATGTGATGGGAACCTATGT	2948
Db	1860	TAAACCCCTCCCATATGTGTGACAGCCTTTCATATGTGAATTAATGTGATGGGAACCTATGT	1919
Qy	2949	GTTACACAGATTAATATCACCAGAAAGCCTGTGATTTCTGAAAAAACTGTGCAGAGCCAAAC	3008
Db	1920	GTTACACAGATTAATATCACCAGAAAGCCTGTGATTTCTGAAAAAACTGTGCAGAGCCAAAC	1979
Qy	3009	TCTGTCAATTTGCAACTCCCACTGTATTTGTATGACAGCACTGTGATATAGTAAGTAATAATTA	3068
Db	1980	TCTGTCAATTTGCAACTCCCACTGTATTTGTATGACAGCACTGTGATATAGTAAGTAATAATTA	2039
Qy	3069	GTACTATTGTGTCAAG 3084	
Db	2040	GTACTATTGTGTCAAG 2055	
RESULT 6			
US-09-780-669-917			
Sequence 917, Application US/09780669			
Patent No. US2002005197A1			
GENERAL INFORMATION:			
APPLICANT: Xu, Jiangchun			
APPLICANT: Dillon, Davin C.			
APPLICANT: Mitcham, Jennifer L.			
APPLICANT: Harlocker, Susan L.			
APPLICANT: Jiang, Yugu			
APPLICANT: Henderson, Robert A.			
APPLICANT: Katos, Michael D.			
APPLICANT: Fanger, Gary R.			
APPLICANT: Retter, Marc W.			
APPLICANT: Stolk, John A.			
APPLICANT: Day, Craig H.			
APPLICANT: Vedvick, Thomas S.			

/ APPLICANT: Carter, Derrick  
 / APPLICANT: Li, Samuel  
 / APPLICANT: Wang, Aijun  
 / APPLICANT: Skeiky, Yasir A.W.  
 / APPLICANT: Hepier, William  
 / APPLICANT: Hural, John  
 / APPLICANT: McNeill, Patricia D.  
 / APPLICANT: Houghton, Raymond L.  
 / TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 / FILE REFERENCE: 210121.427C24  
 / CURRENT APPLICATION NUMBER: US/09/780,669  
 / NUMBER OF SEQ ID NOS: 943  
 / SOFTWARE: FastSeq for Windows Version 3.0  
 / SEQ ID NO 917  
 / LENGTH: 2061  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / US-09-780-669-917

Query Match 64.7%; Score 2028.4; DB 9; Length 2061;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1029 GATTCGACAGCGCATCTTCGACCTTTTCATGTGCGCACACAGCCTTCAGAGCCCTAGGT 1088  
 DB 3 GATTCGACAGCGCATCTTCGACCTTTTCATGTGCGCACACAGCCTTCAGAGCCCTAGGT 62  
 QY 1089 GTCACTGATCAAACTCTTTTCATTCAGAGTCTCTGATTCAGATTTTAATGTTAACT 1148  
 DB 63 GTCACTGATCAAACTCTTTTCATTCAGAGTCTCTGATTCAGATTTTAATGTTAACT 122  
 QY 1149 TTGGAGACAGATATTCAGAAAAAATTTCTTATATAAAAAATACACATCTGATCTTC 1208  
 DB 123 TTGGAGACAGATATTCAGAAAAAATTTCTTATATAAAAAATACACATCTGATCTTC 181  
 QY 1209 AAATATGAACTGGTTGGGGAATTCATTTTTCATATATTTCTTCTTTCTTCTTCT 1268  
 DB 182 AAATATGAACTGGTTGGGGAATTCATTTTTCATATATTTCTTCTTTCTTCTTCT 241  
 QY 1269 TGTACATATATATATTAATACCTGACCTAGTGTGTTGGGGAATTTTCTTCTTCT 1328  
 DB 242 TGTACATATATATTAATACCTGACCTAGTGTGTTGGGGAATTTTCTTCTTCTTCT 301  
 QY 1329 TTTACATGACATCAATCTTAACTGCTTCTTCTGATGTTTACAGCATTTCTGATTA 1388  
 DB 302 TTTACATGACATCAATCTTAACTGCTTCTTCTGATGTTTACAGCATTTCTGATTA 361  
 QY 1389 GAATGATCATCTGAGAGAACTTTGCCAAGCCTTAGACAGCGGAAAGAAATTAACAC 1448  
 DB 362 GAATGATCATCTGAGAGAACTTTGCCAAGCCTTAGACAGCGGAAAGAAATTAACAC 421  
 QY 1449 AGAATATATATAATGAGATATCTAGCTTAAACTTAATCTTCTTCTGAGAACTCCA 1508  
 DB 422 AGAATATATATAATGAGATATCTAGCTTAAACTTAATCTTCTTCTGAGAACTCCA 481  
 QY 1509 ACCACATTTGATCTTCAGAAAAATGCTGTCTTCAAAATGACTCTTACAGAGAAATTAAT 1568  
 DB 482 ACCACATTTGATCTTCAGAAAAATGCTGTCTTCAAAATGACTCTTACAGAGAAATTAAT 541  
 QY 1569 TTTTCTCTGACACTGACATTAAGGGGAAGATGGAATTAAGCCTTGAAGAGGTAC 1628  
 DB 542 TTTTCTCTGACACTGACATTAAGGGGAAGATGGAATTAAGCCTTGAAGAGGTAC 601  
 QY 1629 ATTACCTAGCTTAATGAAGGTGACACATCTTCTGAGATTTTTCACAGCATATGAGACC 1688  
 DB 602 ATTACCTAGCTTAATGAAGGTGACACATCTTCTGAGATTTTTCACAGCATATGAGACC 661  
 QY 1689 CTGTTTCTCATTTAATTTTCTTATCAACCTTTAATTTAGGCAAGATTTATTTAGTAC 1748  
 DB 662 CTGTTTCTCATTTAATTTTCTTATCAACCTTTAATTTAGGCAAGATTTATTTAGTAC 721

QY 1749 CCTCATTTGAGCCATGGGAAAAATGATGTTCACTGGGGATCAGTGAATTTAAATGGGGTCA 1808  
 DB 722 CCTCATTTGAGCCATGGGAAAAATGATGTTCACTGGGGATCAGTGAATTTAAATGGGGTCA 781  
 QY 1809 TACAAGTATTAATAATTAATAAAAAAAGACTTCAATGCCCATCTTCATATGATGAGAG 1868  
 DB 782 TACAAGTATTAATAATTAATAAAAAAAGACTTCAATGCCCATCTTCATATGATGAGAG 839  
 QY 1869 AACTGTTAGAGACCAACAGGGATGAGGTTAGAGATTTCCAGAGCTTATCATTTTCTA 1928  
 DB 840 AACTGTTAGAGACCAACAGGGATGAGGTTAGAGATTTCCAGAGCTTATCATTTTCTA 899  
 QY 1929 GAGAGGTTATTAATTTCTTCTCACTCACTCAGGTGTTGATTTAGAGATTTCTGGCAAC 1988  
 DB 900 GAGAGGTTATTAATTTCTTCTCACTCACTCAGGTGTTGATTTAGAGATTTCTGGCAAC 959  
 QY 1989 AGAAGTCAATGCTTAAATCCCACTAGTATTTGCTATTTGCTGCTCAATGGCAATTA 2048  
 DB 960 AGAAGTCAATGCTTAAATCCCACTAGTATTTGCTATTTGCTGCTCAATGGCAATTA 1019  
 QY 2049 CCTGCTCTTGGAGAGAGATTTCTAGGTTCAACATTAAGAGATTTCTATTCAGAAA 2108  
 DB 1020 CCTGCTCTTGGAGAGAGATTTCTAGGTTCAACATTAAGAGATTTCTATTCAGAAA 1079  
 QY 2109 GTCTGATAGGGCTTATAGCAAGTATTTAATTTTAAAGTTCCATAGGATTTCTGATA 2168  
 DB 1080 GTCTGATAGGGCTTATAGCAAGTATTTAATTTTAAAGTTCCATAGGATTTCTGATA 1139  
 QY 2169 GGCAGTGAAGTTAGGAGACACAGTATATGAGGAATGGAATGAGAGGTCTTGAAG 2228  
 DB 1140 GGCAGTGAAGTTAGGAGACACAGTATATGAGGAATGGAATGAGAGGTCTTGAAG 1199  
 QY 2229 ATTAACATTTGGCTTTTGAAGTGAAGTCTAGTGTGAAGTGAAGGAATCTTCAAGCAAT 2288  
 DB 1200 ATTAACATTTGGCTTTTGAAGTGAAGTCTAGTGTGAAGTGAAGGAATCTTCAAGCAAT 1259  
 QY 2289 GCTTATTTGGGGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2348  
 DB 1260 GCTTATTTGGGGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1319  
 QY 2349 ACTTAGGCAATGGGATCAGGATTTTCTGAGGGCTTATTCAGAGGTTAATGAG 2408  
 DB 1320 ACTTAGGCAATGGGATCAGGATTTTCTGAGGGCTTATTCAGAGGTTAATGAG 1379  
 QY 2409 TTTCACTTCAACAGATATGACAAAGTGTAACTCAAGAACTCAATTAACAATACTA 2468  
 DB 1380 TTTCACTTCAACAGATATGACAAAGTGTAACTCAAGAACTCAATTAACAATACTA 1439  
 QY 2469 AAACATGATCATATATGATGATAGTTTCTTTTCTTCAATCTCAGGTTCCCTGA 2528  
 DB 1440 AAACATGATCATATATGATGATAGTTTCTTTTCTTCAATCTCAGGTTCCCTGA 1499  
 QY 2529 TATGATTTCTTAATAATCATGCTTATCCCTTTTGTATGATGATATTTGGAATGTC 2588  
 DB 1500 TATGATTTCTTAATAATCATGCTTATCCCTTTTGTATGATGATATTTGGAATGTC 1559  
 QY 2589 CTATTTAATCTTGTATTTGCTGCTGACCTGTAAGCCCATGAGGCACTGTTATTTATG 2648  
 DB 1560 CTATTTAATCTTGTATTTGCTGCTGACCTGTAAGCCCATGAGGCACTGTTATTTATG 1619  
 QY 2649 AATGATCTCTGTTCACTGATGACCTGCTTCTGCTATCATTTGAATCCCCAGCAAGT 2708  
 DB 1620 AATGATCTCTGTTCACTGATGACCTGCTTCTGCTATCATTTGAATCCCCAGCAAGT 1679  
 QY 2709 GCTTGAACATATATGATGCTTATGCTTGAACACGGTTATTTTTCATCAAACTGATTTCT 2768  
 DB 1680 GCTTGAACATATATGATGCTTATGCTTGAACACGGTTATTTTTCATCAAACTGATTTCT 1739  
 QY 2769 TCTGCTCTGACATATGACGAGCAATTTTCAGGCTTCTTGAAGTGGGATTTATTTAA 2828  
 DB 1740 TCTGCTCTGACATATGACGAGCAATTTTCAGGCTTCTTGAAGTGGGATTTATTTAA 1799  
 QY 2829 TTTGCGCATTTACTTCAATGATGAGTGAAGTGAATGTCATGTCATTTCTATACCTGCTCA 2888





QY	1989	AGACTCAATGCGCTTAAATCCCAACATGACGATTTGGCTAATGTCCTGGTCCAAATTTGCCAATTA	2048
Db	960	AGAACTATGCGCTTAAATCCCAACATGACGATTTGGCTAATGTCCTGGTCCAAATTTGCCAATTA	1019
QY	2049	CCCTGTGCTTGAGAAAGTGAATTTCTAGATTCCACATTTAGGAAGATTTTATTCAGAAA	2108
Db	1020	CCGTGTGCTTGAGAAAGTGAATTTCTAGATTCCACATTTAGGAAGATTTTATTCAGAAA	1079
QY	2109	GTCTGCATAGGGCTTAATAGCAAGTATTTATTTTAAAGTCCATAGGTATTTCTGATA	2168
Db	1080	GTCTGCATAGGGCTTAATAGCAAGTATTTATTTTAAAGTCCATAGGTATTTCTGATA	1139
QY	2169	GGCATGAGGTTAGGGAGCCCAACGATTTGATGGGAATATGAAATGCGACGCTCTTGAG	2228
Db	1140	GGCATGAGGTTAGGGAGCCCAACGATTTGATGGGAATATGAAATGCGACGCTCTTGAG	1139
QY	2229	ATTAACATTGGCCTTTTGATGTGACTGTAGCTGAAAGTGAAGGAATCTTCAGACCAT	2288
Db	1200	ATTAACATTGGCCTTTTGATGTGACTGTAGCTGAAAGTGAAGGAATCTTCAGACCAT	1259
QY	2289	GCTTATTTTGGGGCTTTTGACAGATATGAAACAGGGACCTTTAGACCAAGGAAGCAATCTG	2348
Db	1260	GCTTATTTTGGGGCTTTTGACAGATATGAAACAGGGACCTTTAGACCAAGGAAGCAATCTG	1319
QY	2349	ACTTAGCGCATGGGAATCAGGCAATTTTGGCTTGAAGGGCTATTACCAAGGTTAATAGG	2408
Db	1320	ACTTAGCGCATGGGAATCAGGCAATTTTGGCTTGAAGGGCTATTACCAAGGTTAATAGG	1379
QY	2409	TTTGCATCTTCAACAGGATATGACAAAGTGTAAACCAAGAAATCTCAATTAACAATACTA	2468
Db	1380	TTTGCATCTTCAACAGGATATGACAAAGTGTAAACCAAGAAATCTCAATTAACAATACTA	1439
QY	2469	AAACATGTGATCATATATGTGTGAAGTTCAATTTCTTTTCAATCCACAGTTCCCTGA	2528
Db	1440	AAACATGTGATCATATATGTGTGAAGTTCAATTTCTTTTCAATCCACAGTTCCCTGA	1439
QY	2529	TATGATTTCTATAACATAGCTTTTCATCCCTTTTGTAAATGATATCATATTTGGAAATGC	2588
Db	1500	TATGATTTCTATAACATAGCTTTTCATCCCTTTTGTAAATGATATCATATTTGGAAATGC	1559
QY	2589	CTATTTAATACCTGATATTTGCTGCGCGACGTGAAGCCCATAGGGCACGTGTTATTTATG	2648
Db	1560	CTATTTAATACCTGATATTTGCTGCGCGACGTGAAGCCCATAGGGCACGTGTTATTTATG	1619
QY	2649	AATGTCAATCTGTTCATCATATGACTGTCTCTTGGCTCATATGAATCCCCAGCAAGT	2708
Db	1620	AATGTCAATCTGTTCATCATATGACTGTCTCTTGGCTCATATGAATCCCCAGCAAGT	1679
QY	2709	GCCATGAACATATAAGTGCCTAATGCTTGAACACCGGTTATTTTGCATCAAACTGATTCCT	2768
Db	1680	GCCATGAACATATAAGTGCCTAATGCTTGAACACCGGTTATTTTGCATCAAACTGATTCCT	1739
QY	2769	TCTGTGCTGGAACAGATAGCCAGGCAATTTTCCAGCTTCTTTGATGGGTGATTTATTTAA	2828
Db	1740	TCTGTGCTGGAACAGATAGCCAGGCAATTTTCCAGCTTCTTTGATGGGTGATTTATTTAA	1739
QY	2829	TTCTGGGCATTAATCTTCCAAATGTGATGGAATGACATGTGCAATTTCTATACCTGGCTCA	2888
Db	1800	TTCTGGGCATTAATCTTCCAAATGTGATGGAATGACATGTGCAATTTCTATACCTGGCTCA	1859
QY	2889	TAAAAACCTCCCATGTGACGCTTTCATAGTTGAACATTTAAATGTACCTGGGAACCTATGT	2948
Db	1860	TAAAAACCTCCCATGTGACGCTTTCATAGTTGAACATTTAAATGTACCTGGGAACCTATGT	1919
QY	2949	GTATACACAGATTAATCACCAAGACCTGTGATTTCTGAAAAAACTGTGCAGAGCCAAAC	3008
Db	1920	GTATACACAGATTAATCACCAAGAGCCGTGATTTCTGAAAAAACTGTGCAGAGCCAAAC	1979
QY	3009	TCTGTCAATTTGCAACTCCCATTTGATTTTGAACGAGGACGTTGATTAAGTAAAAAATPAA	3068
Db	1980	TCTGTCAATTTGCAACTCCCATTTGATTTTGAACGAGGACGTTGATTAAGTAAAAAATPAA	2039
QY	3069	GTACTATTTGTGTCAAG	3084

[illegible]

Db 362 GAATGTCATCTTAAAGAACTTTGGCAAGGCTTAAAGGCAAGGAAATTAACAC 421  
 Qy 1449 AGAATTAATAAATAAGATTAATCTAGCTTAAATCTATCTCTCTTCAAGATCCCA 1508  
 Db 442 AGAATTAATAAATAAGATTAATCTAGCTTAAATCTATCTCTCTTCAAGATCCCA 481  
 Qy 1509 ACCAATAGATCTGAGAAATGCTCTTCAAAATGACTTTCACAGAGAAATTAAT 1568  
 Db 482 ACCAATAGATCTGAGAAATGCTCTTCAAAATGACTTTCACAGAGAAATTAAT 541  
 Qy 1569 TTTTCCTGCACTAGCACTTAAAGGGAAGATTGGAAGTAAAGCTTGAAGAGTAC 1628  
 Db 542 TTTTCCTGCACTAGCACTTAAAGGGAAGATTGGAAGTAAAGCTTGAAGAGTAC 601  
 Qy 1629 ATTTCCTAGCTTAATGAAAGTGAACAACCTGCTTGAAGTTCACAGATATGAGAC 1688  
 Db 602 ATTTCCTAGCTTAATGAAAGTGAACAACCTGCTTGAAGTTCACAGATATGAGAC 661  
 Qy 1689 CTGTTTCTCTATTAATTTTCTTATCAACCTTAAATAGGCAAGATTAATATGATAC 1748  
 Db 662 CTGTTTCTCTATTAATTTTCTTATCAACCTTAAATAGGCAAGATTAATATGATAC 721  
 Qy 1749 CCTCATTTAGCACTGGAATAATGATGCTGAGTGGGATCAAGTAAATGAGGCTCA 1808  
 Db 722 CCTCATTTAGCACTGGAATAATGATGCTGAGTGGGATCAAGTAAATGAGGCTCA 781  
 Qy 1809 TACAAGTATTAATAATTAATAAATAAAGACTTCATGCTCCCACTTCATATGATGAGAG 1868  
 Db 782 TACAAGTATTAATAATTAATAAATAAAGACTTCATGCTCCCACTTCATATGATGAGAG 839  
 Qy 1869 AACTGTTAGAGAACCAACAGGGTAGTGGGTTAGAGATTCACAGATCTTACATTTTCTA 1928  
 Db 840 AACTGTTAGAGAACCAACAGGGTAGTGGGTTAGAGATTCACAGATCTTACATTTTCTA 899  
 Qy 1929 GAGAGGATTAATTTTCTTCTCACTCATCCAGTGTGATTTAGAAATTTCTGAGAAC 1988  
 Db 900 GAGAGGATTAATTTTCTTCTCACTCATCCAGTGTGATTTAGAAATTTCTGAGAAC 959  
 Qy 1989 AGAATCTAGGCTTAAATCCCACTAGATTTGCTTATGCTGCTGCTCAATGCAATTA 2048  
 Db 960 AGAATCTAGGCTTAAATCCCACTAGATTTGCTTATGCTGCTGCTCAATGCAATTA 1019  
 Qy 2049 CCTGCTCTGGAAGAAATGTTCTAGCTTCACTATTAAGATTTCTTATTCAGAA 2108  
 Db 1020 CCTGCTCTGGAAGAAATGTTCTAGCTTCACTATTAAGATTTCTTATTCAGAA 1079  
 Qy 2109 GTCGCAATAGGCTTATAGCAAGTATTTATTTAAAGTCCATAGGATTCGATA 2168  
 Db 1080 GTCGCAATAGGCTTATAGCAAGTATTTATTTAAAGTCCATAGGATTCGATA 1139  
 Qy 2169 GGCAGTGAAGTTAGGAGCCACAGTTATGATGGAAGTGAATGGCAGGCTTGAAG 2228  
 Db 1140 GGCAGTGAAGTTAGGAGCCACAGTTATGATGGAAGTGAATGGCAGGCTTGAAG 1199  
 Qy 2229 ATACATTTGAGCTTTTATGATGCTGCTAGCTGGAAGAGGGAATCTTCAAGCAAT 2288  
 Db 1200 ATACATTTGAGCTTTTATGATGCTGCTAGCTGGAAGAGGGAATCTTCAAGCAAT 1259  
 Qy 2289 GCTTATTTGGGGCTTTGTCAGATAGAAACAGGACTTTGAGACCAAGAAACCAATCTG 2348  
 Db 1260 GCTTATTTGGGGCTTTGTCAGATAGAAACAGGACTTTGAGACCAAGAAACCAATCTG 1319  
 Qy 2349 ACTTAGGCAATGGAATAGGCAATTTTGTCTGAGAGGGCTATTAACAGGGTATATAG 2408  
 Db 1320 ACTTAGGCAATGGAATAGGCAATTTTGTCTGAGAGGGCTATTAACAGGGTATATAG 1379  
 Qy 2409 TTTTATTTTCAACAGATATGACAAACAGTGTAAACCAAGAACTCAATTTACAAATATCA 2468  
 Db 1380 TTTTATTTTCAACAGATATGACAAACAGTGTAAACCAAGAACTCAATTTACAAATATCA 1439  
 Qy 2469 AAACATGATCATATATGATGATGATTTTCTTTTCAATCTCAGGATCCCTCA 2528

Db 1440 AAACATGATCATATATGATGATGATTTTCTTTTCAATCTCAGGATCCCTCA 1499  
 Qy 2529 TATGATCTCTAATACATGCTTTCATCCCTTTTGAATGATATCATATTTGAAATGC 2588  
 Db 1500 TATGATCTCTAATACATGCTTTCATCCCTTTTGAATGATATCATATTTGAAATGC 1559  
 Qy 2589 CTATTTAATATCTGATTTTGTCTGAGACTGTAAGCCATGAGGCACTGTTATATG 2648  
 Db 1560 CTATTTAATATCTGATTTTGTCTGAGACTGTAAGCCATGAGGCACTGTTATATG 1619  
 Qy 2649 AATGATCTCTGATCATATGATGCTCTTGTCTGATGATGATGATGATGATGATGAT 2708  
 Db 1620 AATGATCTCTGATCATATGATGCTCTTGTCTGATGATGATGATGATGATGATGAT 1679  
 Qy 2709 GCTTAGAATTAATAGTGTATAGCTTGAACACCGGTTATTTTCAACACCTGATCT 2768  
 Db 1680 GCTTAGAATTAATAGTGTATAGCTTGAACACCGGTTATTTTCAACACCTGATCT 1739  
 Qy 2769 TCTGCTGGAACACATAGCAAGGCAATTTTCAAGCTTGTGATGATGATGATGATGAT 2828  
 Db 1740 TCTGCTGGAACACATAGCAAGGCAATTTTCAAGCTTGTGATGATGATGATGATGAT 1799  
 Qy 2829 TTTGCGCATTAATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2888  
 Db 1800 TTTGCGCATTAATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1859  
 Qy 2889 TAAAACCTTCCATGCTGAGCTTTCATGTTGACATTAATGATGATGATGATGATGAT 2948  
 Db 1860 TAAAACCTTCCATGCTGAGCTTTCATGTTGACATTAATGATGATGATGATGATGAT 1919  
 Qy 2949 GTTACACAGATTAATCAACCAAGGCTGATTTTGTGAATAAAGCTGCAAGCAAC 3008  
 Db 1920 GTTACACAGATTAATCAACCAAGGCTGATTTTGTGAATAAAGCTGCAAGCAAC 1979  
 Qy 3009 TCTGCTATTTGCACTCCCACTGTTGATTTGATGAGGAGTGTGATTAAGTAAATTA 3068  
 Db 1980 TCTGCTATTTGCACTCCCACTGTTGATTTGATGAGGAGTGTGATTAAGTAAATTA 2039  
 Qy 3069 GTACTATTTGTCAAG 3084  
 Db 2040 GTACTATTTGTCAAG 2055

RESULT 10  
 US-10-144-678A-917  
 ; Sequence 917, Application US/10144678A  
 ; Publication No. US20030157089A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Dillon, David C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Panger, Gary R.  
 ; APPLICANT: Retter, Marc W.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Vedrick, Thomas S.  
 ; APPLICANT: Carter, Darick  
 ; APPLICANT: Li, Samuel X.  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Skeiky, Yasir A. W.  
 ; APPLICANT: Hepler, William T.  
 ; APPLICANT: Hurst, John  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Houghton, Raymond L.  
 ; APPLICANT: Vinals y de Bassole, Carlota  
 ; APPLICANT: Foy, Teresa M.  
 ; APPLICANT: Matanabe, Yoshihiro  
 ; APPLICANT: Deng, Ta  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

```

1  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
2
3  FILE REFERENCE:  210121.427C28
4
5  CURRENT APPLICATION NUMBER:  US/10/144,678A
6
7  CURRENT FILING DATE:  2002-08-12
8
9  NUMBER OF SEQ ID NOS:  1033
10
11  SOFTWARE:  FASTSEQ for Windows Version 3.0
12
13  SEQ ID NO 917
14
15  LENGTH:  2061
16
17  TYPE:  DNA
18
19  ORGANISM:  Homo sapiens
20
21  US-10-144-678A-917

```

Query Match	64.7%	Score 2028.4	DB 13	Length 2061
Best Local Similarity	99.8%	Pred. No. 0		
Matches 2052; Conservative	0	Mismatches	1	Indels 3; Gaps 2

[illegible]

Db	840	AACCTTTAGAGACCAACAGGGTGTGGTTAGAGATTTCCAGATCTTACATTTCTTA	899
Qy	1929	GAGGAGTATTTAATTTCTTCTCACTCACTCAAGTGTGATTTAGAAATTTCTGGCAAC	1988
Db	900	GAGGAGTATTTAATTTCTTCTCACTCACTCAAGTGTGATTTAGGAATTTCTGGCAAC	959
Qy	1989	AGAACTCAAGCTTAAATCCCACTAGCAATGTCTATGTCTGTGTCATATGCCAATTA	204
Db	960	AGAACTCAAGCTTAAATCCCACTAGCAATGTCTATGTCTGTGTCATATGCCAATTA	1015
Qy	2049	CCGTGTCTTGGAAAGATGAATTTCTAGGTTACCATATGGAAGATCTTATTCAGAAA	2108
Db	1020	CCGTGTCTTGGAAAGATGAATTTCTAGGTTACCATATGGAAGATCTTATTCAGAAA	1079
Qy	2109	GTCTGCATAGGGCTTATAGCAAGTATTTATTTTAAAAGTTCCATAGGTATCTGATA	2166
Db	1080	GTCTGCATAGGGCTTATAGCAAGTATTTATTTTAAAAGTTCCATAGGTATCTGATA	1139
Qy	2169	GGCAGTAGGTATTAGGGAGCCACAGTTATGATGGGAAGATATGGAATGGCAGGTCTTGAAG	2222
Db	1140	GGCAGTAGGTATTAGGGAGCCACAGTTATGATGGGAAGATATGGAATGGCAGGTCTTGAAG	1199
Qy	2229	ATTAACATTGSCCTTTTGAATGTGACTGTAGCTGTGAAAAGTAGGGAAATCTTCAGACCAT	2288
Db	1200	ATTAACATTGSCCTTTTGAATGTGACTGTAGCTGTGAAAAGTAGGGAAATCTTCAGACCAT	1259
Qy	2289	GCTTATTTGGGGCTTTGTGCAGATATGGAAACAGGAATTTTGAAGCCAGGAAGCAATCTG	2344
Db	1260	GCTTATTTGGGGCTTTGTGCAGATATGGAAACAGGAATTTTGAAGCCAGGAAGCAATCTG	1311
Qy	2349	ACTTAGGCATGGGAATCAAGCATTTTGTCTGTAGGGGCTTATACCAAGGCTATATAG	2408
Db	1320	ACTTAGGCATGGGAATCAAGCATTTTGTCTGTAGGGGCTTATACCAAGGCTATATAG	1379
Qy	2409	TTTCATCTTTCAACAGGATATGACAAACAGTGTAAACAAAGAACTCAATTAACAATACTA	2466
Db	1380	TTTCATCTTTCAACAGGATATGACAAACAGTGTAAACAAAGAACTCAATTAACAATACTA	1433
Qy	2469	AAACATGTGATCATATATGTGTGAAGTTCAATTTCTTTCAATCCCAAGTTCCTCTGA	2528
Db	1440	AAACATGTGATCATATATGTGTGAAGTTCAATTTCTTTCAATCCCAAGTTCCTCTGA	1499
Qy	2529	TATGATTTCCATTAACATGCTTTATCCCTTTTGTAATGGAATCATATTTGGAAATGC	2588
Db	1500	TATGATTTCCATTAACATGCTTTATCCCTTTTGTAATGGAATCATATTTGGAAATGC	1559
Qy	2589	CTATTTAATACTTGTATTTGTGTGTGACGTGAAGCCATGAGGCACTGTGTAATTTATG	2648
Db	1560	CTATTTAATACTTGTATTTGTGTGTGACGTGAAGCCATGAGGCACTGTGTAATTTATG	1618
Qy	2649	AATGTCAATCTCTGTTCATCATATGACTGTCTTTTGCTCATCATATGAAATCCCCAGCAAGT	2708
Db	1620	AATGTCAATCTCTGTTCATCATATGACTGTCTCTTTTGCTCATCATATGAAATCCCCAGCAAGT	1679
Qy	2709	GCCATGAAGCAATATAGTGTATATGCTTGACACCGGTTATTTTATATCAAAACCTGATTCCT	2768
Db	1680	GCCATGAAGCAATATAGTGTATATGCTTGACACCGGTTATTTTATATCAAAACCTGATTCCT	1739
Qy	2769	TCTGTCTGTGAACATATAGCAAGGCAATTTTTCAGACCTCTTTTGAGTTGGGTATTTAAA	2828
Db	1740	TCTGTCTGTGAACATATAGCAAGGCAATTTTTCAGACCTCTTTTGAGTTGGGTATTTAAA	1799
Qy	2829	TTCTGGCCATTAATCTCCAAATGTAGTGGAAAGTACATGTGCAATTTCTATATCTGGCTCA	2888
Db	1800	TTCTGGCCATTAATCTCCAAATGTAGTGGAAAGTACATGTGCAATTTCTATATCTGGCTCA	1859
Qy	2889	TAAAAACCTCCCATATGTGCAGACCTTTCAATGTGACATTAATAATGTACATTGGGAACCTATGT	2948
Db	1860	TAAAAACCTCCCATATGTGCAGACCTTTCAATGTGACATTAATAATGTACATTGGGAACCTATGT	1919
Qy	2949	GTTTACACAGATTAATACCAAGAGCTGTGATTTCTGAAAAAACTGTGCAGAGCCAAACC	3008



Db 1920 GTTACACAGATTAATACACAGAGCTGATTTCTGAAAAAAGTGCAGAGCCAAACC 1979  
QY 3009 TCTGTCATTTGGAACTCCCACTTGTATTTGTACGAGGAGATTGATGAAAAATTA 3068  
Db 1980 TCTGTCATTTGGAACTCCCACTTGTATTTGTACGAGGAGATTGATGAAAAATTA 2039  
QY 3069 GTACTATTTGTGTAAG 3084  
Db 2040 GTACTATTTGTGTAAG 2055

RESULT 11  
US-10-294-025-917  
Sequence 917, Application US/10294025  
Publication No. US20030185830A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jianshun  
APPLICANT: Stolk, John A.  
APPLICANT: Kalos, Michael D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C29  
CURRENT APPLICATION NUMBER: US/10/294,025  
NUMBER OF SEQ ID NOS: 1038  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 917  
LENGTH: 2061  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-294-025-917

Query Match 64.7%; Score 2028.4; DB 13; Length 2061;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1029 GATTGACAGGCGATCCTTCGACTTTTCATGTGCGCCACAGCCCTCAGAGCCCTAGT 1088  
Db 3 GATTGACAGGCGATCCTTCGACTTTTCATGTGCGCCACAGCCCTCAGAGCCCTAGT 62  
QY 1089 GTGAGTATTCAACTCTTTTCCATTCAGAGTCTCTGATTCAGATTTTAATGTTAAT 1148  
Db 63 GTGAGTATTCAACTCTTTTCCATTCAGAGTCTCTGATTCAGATTTTAATGTTAAT 122  
QY 1149 TTTGGAGACAGTATTCAGAAAAAATTTCTTAATAAAAATACAACTCAGATCCTTC 1208  
Db 123 TTTGGAGACAGTATTCAGAAAAAATTTCTTAATAAAAATACAACTCAGATCCTTC 181  
QY 1209 AAATATGAACCTGTTGGGGAATCTCCATTTTTCATTAATTTTCTTTGTTTCT 1268  
Db 182 AAATATGAACCTGTTGGGGAATCTCCATTTTTCATTAATTTTCTTTGTTTCT 241  
QY 1269 TGCTACATATTAATTAATACCTGACTAGTGTGTTGGAGGTTATTAATTTTCT 1328  
Db 242 TGCTACATATTAATTAATACCTGACTAGTGTGTTGGAGGTTATTAATTTTCT 301  
QY 1329 TTTACCATGAGTCCAAATCTAACTGTTTCTACTGATGTTTACACATTTCTGAGATA 1388  
Db 302 TTTACCATGAGTCCAAATCTAACTGTTTCTACTGATGTTTACACATTTCTGAGATA 361  
QY 1389 GAATGTCATCTGAGAGAACTTTGCCAAGGCTTAAGCAAGGGAAGAAAAATTAAC 1448  
Db 362 GAATGTCATCTGAGAGAACTTTGCCAAGGCTTAAGCAAGGGAAGAAAAATTAAC 421  
QY 1449 AGAATATATATAATGAGATATCTAGCTTAAATCTAACTTCTCTTCTGAGAACTCCCA 1508  
Db 422 AGAATATATATAATGAGATATCTAGCTTAAATCTAACTTCTCTTCTGAGAACTCCCA 481  
QY 1509 ACCCATTTGATCTCAGAAAAATGCTGTTCAAAAATGACTTTACAGAGAAAGAAATTAAT 1568  
Db 482 ACCCATTTGATCTCAGAAAAATGCTGTTCAAAAATGACTTTACAGAGAAAGAAATTAAT 541  
QY 1569 TTTTCTCTGAGACACTAGCACTTAAGGGGAAGATTGAGATAAAGCTTTGAAAGATAC 1628

Db 542 TTTTCTCTGAGACACTAGCACTTAAGGGGAAGATTGAGATAAAGCTTTGAAAGATAC 601  
QY 1629 ATTACCTTAGTAAATGAAAGTTGACACACTGTTCTGAGAGTTTTCACAGATATGAGC 1688  
Db 602 ATTACCTTAGTAAATGAAAGTTGACACACTGTTCTGAGAGTTTTCACAGATATGAGC 661  
QY 1689 CTGTTTCTCTGATTAATTTTCTTATCAACCTTTTAATTAAGGCAAGATATTAATGATC 1748  
Db 662 CTGTTTCTCTGATTAATTTTCTTATCAACCTTTTAATTAAGGCAAGATATTAATGATC 721  
QY 1749 CCTCATTTGAGCCATGAGAAAAATGATGTTCAAGTGGGATCAAGTAATTAATGAGGATCA 1808  
Db 722 CCTCATTTGAGCCATGAGAAAAATGATGTTCAAGTGGGATCAAGTAATTAATGAGGATCA 781  
QY 1809 TACAATATTAATAATTAATAAAGAAAAAGCTTCATGCCCATCTCATATGATGAGAG 1868  
Db 782 TACAATATTAATAATTAATAAAGAAAAAGCTTCATGCCCATCTCATATGATGAGAG 839  
QY 1869 AACTGTTAGAGACCAACAGGGTAGTGGGTTAGAGATTCCAGAGTCTTAATTTCTA 1928  
Db 840 AACTGTTAGAGACCAACAGGGTAGTGGGTTAGAGATTCCAGAGTCTTAATTTCTA 899  
QY 1929 GAGAGGATTTTAATTTCTTCACTCATCCAGTGTGATTTAGAAATTTCTGCGAAC 1988  
Db 900 GAGAGGATTTTAATTTCTTCACTCATCCAGTGTGATTTAGAAATTTCTGCGAAC 959  
QY 1989 AGAAGTATGAGCTTAATCCCACTAGCTATGCTTATGTTCTGTTCCATGCAATTA 2048  
Db 960 AGAAGTATGAGCTTAATCCCACTAGCTATGCTTATGTTCTGTTCCATGCAATTA 1019  
QY 2049 CCTGTCTTGGAGAGAGATTTTCTAGGTTCAACATTAAGAAATTTCTTATCAGAAA 2108  
Db 1020 CCTGTCTTGGAGAGAGATTTTCTAGGTTCAACATTAAGAAATTTCTTATCAGAAA 1079  
QY 2109 GTGTGATAGGCTTAATGCAAGTATTAATTTTAAAGTCCATAGTGAATTTGATA 2168  
Db 1080 GTGTGATAGGCTTAATGCAAGTATTAATTTTAAAGTCCATAGTGAATTTGATA 1139  
QY 2169 GGCAGTATGAGGATGAGAGCAACAGTATGATGAGAAATGAGATGCAAGGCTTTGAAG 2228  
Db 1140 GGCAGTATGAGGATGAGAGCAACAGTATGATGAGAAATGAGATGCAAGGCTTTGAAG 1199  
QY 2229 ATTAATGAGGCTTTGAGTGAATCGTATGAGTGAAGGAAATGAGGAACTTCAAGCAAT 2288  
Db 1200 ATTAATGAGGCTTTGAGTGAATCGTATGAGTGAAGGAAATGAGGAACTTCAAGCAAT 1259  
QY 2289 GCTTATTTGGGCTTTGAGATGAGAACAGGAACTTTGAGACAGAGAAAGCAATCTG 2348  
Db 1260 GCTTATTTGGGCTTTGAGATGAGAACAGGAACTTTGAGACAGAGAAAGCAATCTG 1319  
QY 2349 ACTTGGAGATGAGGAAATCAGGCACTTTTCTCTGAGGGGCTATTAACAGGTTAAATAG 2408  
Db 1320 ACTTGGAGATGAGGAAATCAGGCACTTTTCTCTGAGGGGCTATTAACAGGTTAAATAG 1379  
QY 2409 TTTCACTTTCACACAGATATGCAACAGTGTAAACCAAGAACTCAATTAATACTA 2468  
Db 1380 TTTCACTTTCACACAGATATGCAACAGTGTAAACCAAGAACTCAATTAATACTA 1439  
QY 2469 AAACATGATATATATGATGTAAGTTTCACTTTTCTTCAATCTCAAGTTCCCTGA 2528  
Db 1440 AAACATGATATATATGATGTAAGTTTCACTTTTCTTCAATCTCAAGTTCCCTGA 1499  
QY 2529 TATGATTTCTTAATAATCATGCTTTCATCCCTTTTGTAAATGATATCATATTTGAAATGC 2588  
Db 1500 TATGATTTCTTAATAATCATGCTTTCATCCCTTTTGTAAATGATATCATATTTGAAATGC 1559  
QY 2589 CTATTTATATCTGATTTGCTGCTGCACTGAAGCCATGAGGCACTGTTTATATTTG 2648  
Db 1560 CTATTTATATCTGATTTGCTGCTGCACTGAAGCCATGAGGCACTGTTTATATTTG 1619  
QY 2649 AATGTCATCTCTGTCATCATTTGACTGCTTTTCTCTCATTTGAATCCCCAGCAAGT 2708



Db 1620 AATGCTCTCTGTTCAATGATGAGCTCTTTCATCATGAAATCCCGAGAAAGT 1679  
Qy 2709 GCGTAGAACAATATAGTCTTATGCTTACAGACCGGTTATTTTCATCAAACTGATTCCT 2768  
Db 1680 GCGTAGAACAATATAGTCTTATGCTTACAGACCGGTTATTTTCATCAAACTGATTCCT 1739  
Qy 2769 TCTGCTCTGAAACATATGCGGCAATTTTTCAGCTCTTCTTGAAGTTATTTAA 2828  
Db 1740 TCTGCTCTGAAACATATGCGGCAATTTTTCAGCTCTTCTTGAAGTTATTTAA 1799  
Qy 2829 TTGCGCATTAATCTTCAATGAGTGAAGTGAATGCAATTTCAATACCTGAGCTA 2888  
Db 1800 TTGCGCATTAATCTTCAATGAGTGAAGTGAATGCAATTTCAATACCTGAGCTA 1859  
Qy 2889 TAAACCTCTCCATGTCAGCTTTCATGTTGAATTAATGAGCTTGGGAAGCTATGT 2948  
Db 1860 TAAACCTCTCCATGTCAGCTTTCATGTTGAATTAATGAGCTTGGGAAGCTATGT 1919  
Qy 2949 GTTACAGAGATTAATACAGAGAGCTGATTTCTGAAAACCTGCGAGAGCCAAAC 3008  
Db 1920 GTTACAGAGATTAATACAGAGAGCTGATTTCTGAAAACCTGCGAGAGCCAAAC 1979  
Qy 3009 TCTGTCAATTTGCACTCCCACTTGTATTTGTACAGAGCACTTGATAGTGAATAATA 3068  
Db 1980 TCTGTCAATTTGCACTCCCACTTGTATTTGTACAGAGCACTTGATAGTGAATAATA 2039  
Qy 3069 GTACTATTGTGTCAAG 3084  
Db 2040 GTACTATTGTGTCAAG 2055

## RESULT 12

US-10-012-896-917  
Sequence 917, Application US/10012896  
Publication No. US20020183251A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yugu  
APPLICANT: Kalos, Michael D.  
APPLICANT: Reiter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Derrick  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Hurai, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vitale de Bassols, Carolea  
APPLICANT: Foy, Teresa  
APPLICANT: Fanger, Gary R.  
APPLICANT: Mantanabe, Yoshihiro  
APPLICANT: Meagher, Madeleine Joy  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427027  
CURRENT APPLICATION NUMBER: US/10/012,896  
CURRENT FILING DATE: 2001-12-10  
NUMBER OF SEQ ID NOS: 1011  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 917  
LENGTH: 2061  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-012-896-917

Query Match 64.7%; Score 2028.4; DB 14; Length 2061;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;  
Qy 1029 GATTGACAGAGGACCTCTGACATTTTCATATGAGCCACACAGCTTCAGAGCCCTAGGT 1088  
Db 3 GATTGACAGAGGACCTCTGACATTTTCATATGAGCCACACAGCTTCAGAGCCCTAGGT 62  
Qy 1089 GTGAGTATCAAACTCTTTCATATGAGTCTGATTCAGATTTTAATGTTAACT 1148  
Db 63 GTGAGTATCAAACTCTTTCATATGAGTCTGATTCAGATTTTAATGTTAACT 122  
Qy 1149 TTGGAGACAGTATTCAGAAAAAATTTCTTAATAAAAATACACCTCAGATCCTTC 1208  
Db 123 TTGGAGACAGTATTCAGAAAAAATTTCTTAATAAAAATACACCTCAGATCCTTC 181  
Qy 1209 AAATATGAACTGGTGGGGAATCCATTTTCAATATATTTCTGCTGTTTCT 1268  
Db 182 AAATATGAACTGGTGGGGAATCCATTTTCAATATATTTCTGCTGTTTCT 241  
Qy 1269 TGTACATATATATTAATACCTGATAGTGTGTGGAGGTTATTAATCTTTTCAT 1328  
Db 242 TGTACATATATATTAATACCTGATAGTGTGTGGAGGTTATTAATCTTTTCAT 301  
Qy 1329 TTACCATGACGTCCAACTTAACCTGCTTACTGATGTTTACAGATTTGAGATTA 1388  
Db 302 TTACCATGACGTCCAACTTAACCTGCTTACTGATGTTTACAGATTTGAGATTA 361  
Qy 1389 GAATGTATCATCTAGAGAACATTTGCCAAGGCTTAAGCAGGCAAGAAATTAACAC 1448  
Db 362 GAATGTATCATCTAGAGAACATTTGCCAAGGCTTAAGCAGGCAAGAAATTAACAC 421  
Qy 1449 AGAATATATATATATGATGATTAATGATTAATGATTAATGATTAATGATTAAT 1508  
Db 422 AGAATATATATATATGATGATTAATGATTAATGATTAATGATTAATGATTAAT 481  
Qy 1509 ACCCATTTGATTCAGAAAAATGCTGCTCAAAAGATGCTTACAGAGAGAAATAT 1568  
Db 482 ACCCATTTGATTCAGAAAAATGCTGCTCAAAAGATGCTTACAGAGAGAAATAT 541  
Qy 1569 TTTTCTCTGACACTAGCACTTAAGGGAAGATTTGGAAGTAAAGCTTGAAGAGTAC 1628  
Db 542 TTTTCTCTGACACTAGCACTTAAGGGAAGATTTGGAAGTAAAGCTTGAAGAGTAC 601  
Qy 1629 ATTTACCTAGTATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1688  
Db 602 ATTTACCTAGTATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 661  
Qy 1689 CTGTTTTTCTATTTAATTTTCTATCAACCTTTAATTAAGCAAGATATTTAATGATTA 1748  
Db 662 CTGTTTTTCTATTTAATTTTCTATCAACCTTTAATTAAGCAAGATATTTAATGATTA 721  
Qy 1749 CCTATTGTAGCACTGGAATAATGATGTTCACTGGGATCACTGAATTAATGAGGCTCA 1808  
Db 722 CCTATTGTAGCACTGGAATAATGATGTTCACTGGGATCACTGAATTAATGAGGCTCA 781  
Qy 1809 TACAAGTATAAAAATTAATAAAAAAAGACTCATGCCAATCTCATATGATGAGAG 1868  
Db 782 TACAAGTATAAAAATTTAAAAAAGACTCATGCCAATCTCATATGATGAGAG 839  
Qy 1869 AACTGTTAGAGACCAAGGAGTGGGTGAGATTTCCAGAGTCTTAATTTCTA 1928  
Db 840 AACTGTTAGAGACCAAGGAGTGGGTGAGATTTCCAGAGTCTTAATTTCTA 899  
Qy 1929 GAGGAGTATTTAATTTCTTCACTCATCAAGTGTGTTAAGAAATTTCTGCGAAC 1988  
Db 900 GAGGAGTATTTAATTTCTTCACTCATCAAGTGTGTTAAGAAATTTCTGCGAAC 959  
Qy 1989 AGAACTCATGCTTAATTAATCCACTAGCTATGTTATGCTGCTCAATTTCCAAATTA 2048  
Db 960 AGAACTCATGCTTAATTAATCCACTAGCTATGTTATGCTGCTCAATTTCCAAATTA 1019  
Qy 2049 CCGTGTCTTGAAGAAAGTATTTCTAGTTCAACATTAAGAAAGTCTTATTCAGAAA 2108

Db	1020	CTGTGCTTGGAAAGAGTATTCCTAGGTCACCATTAATGGAAAGATTCCTATTCAGAA	1079
QY	2109	GTCGSCATAGGGCCTTAATGCAAGTATATTAATTTTAAAGTCCATAGTGAATCTGATA	2168
Db	1080	GTCGSCATAGGGCCTTAATGCAAGTATATTAATTTTAAAGTCCATAGTGAATCTGATA	1139
QY	2169	GGCAGTGAAGTTAGGAGCCACAGTTATGATGGAGATGGAATGGCAGGCTTTGAAG	2228
Db	1140	GGCAGTGAAGTTAGGAGCCACAGTTATGATGGAGATGGAATGGCAGGCTTTGAAG	1139
QY	2229	ATPAATTTGGCCTTTTGAATGGAATCTGAGAGTGGAAAGTGAAGGAATCTTCAGACAT	2288
Db	1200	ATPAATTTGGCCTTTTGAATGGAATCTGAGAGTGGAAAGTGAAGGAATCTTCAGACAT	1259
QY	2289	GCTTAATTTGGGCTTTGTGCAATGGAACAAGGACTTTGAGACACAGAAAGCAATCTG	2348
Db	1260	GCTTAATTTGGGCTTTGTGCAATGGAACAAGGACTTTGAGACACAGAAAGCAATCTG	1319
QY	2349	ACTTAGGATGGGAATTCAGGCAATTTTGGCTCTGAGGGGCTAATACCAAGGCTTAATAG	2408
Db	1320	ACTTAGGATGGGAATTCAGGCAATTTTGGCTCTGAGGGGCTAATACCAAGGCTTAATAG	1379
QY	2409	TTTCACTTTCAACAGGATATGACAAACAGTGTAAACCAAGAACTCAATTAACAATACTA	2468
Db	1380	TTTCACTTTCAACAGGATATGACAAACAGTGTAAACCAAGAACTCAATTAACAATACTA	1439
QY	2469	AAACATGATCATAATATGAGTAAAGTTCAATTTCTTTTCATTCCTCAGGTCCTCGA	2528
Db	1440	AAACATGATCATAATATGAGTAAAGTTCAATTTCTTTTCATTCCTCAGGTCCTCGA	1499
QY	2529	TATGATTTCCATAAATCAGCTTTTCAATCCCTTTTGTAAATGATATCATATTTGGAAATGC	2588
Db	1500	TATGATTTCCATAAATCAGCTTTTCAATCCCTTTTGTAAATGATATCATATTTGGAAATGC	1559
QY	2589	CTATTTAATCTTGATATTTGCTGCGAGCTTAAGCCATGAGGGCACTGTTATTTATG	2648
Db	1560	CTATTTAATCTTGATATTTGCTGCGAGCTTAAGCCATGAGGGCACTGTTATTTATG	1619
QY	2649	AATGCACTCTGTTTCATCATTTGACTGCTCTTTGCTCATCATTAATCCCCAGAAAGT	2708
Db	1620	AATGCACTCTGTTTCATCATTTGACTGCTCTTTGCTCATCATTAATCCCCAGAAAGT	1679
QY	2709	GCCTAGAACAATAATAGTCTTAATCTTGACACCCGTTATTTTTCATCAAACTGATTCCT	2768
Db	1680	GCCTAGAACAATAATAGTCTTAATCTTGACACCCGTTATTTTTCATCAAACTGATTCCT	1739
QY	2769	TCTGTCCCTGGAACAATAGCCAGCAATTTTCCACCCCTCTTTGATGGGTAATATTTAA	2828
Db	1740	TCTGTCCCTGGAACAATAGCCAGCAATTTTCCACCCCTCTTTGATGGGTAATATTTAA	1799
QY	2829	TTCTGGCCATTACTTCCATATGTAATGTAAGTGAACATGTGCATTTCTAATACCTGCTCA	2888
Db	1800	TTCTGGCCATTACTTCCATATGTAATGTAAGTGAACATGTGCATTTCTAATACCTGCTCA	1859
QY	2889	TAAACACCTCCCATATGATGAGCCTTCAATGTTGAATTAATGTAATTTGGGAAGCTATGT	2948
Db	1860	TAAACACCTCCCATATGATGAGCCTTCAATGTTGAATTAATGTAATTTGGGAAGCTATGT	1919
QY	2949	GTTACACAGATAATCAACAGAAAGCTGTGATTTCTGAAAAAACTGTGCAGAGCCAAACC	3008
Db	1920	GTTACACAGATAATCAACAGAAAGCTGTGATTTCTGAAAAAACTGTGCAGAGCCAAACC	1979
QY	3009	TCTGTCAATTTGCAACTCCCACTTGTATTTTGAACAGGCAATGGAATAAGAAAAATAA	3068
Db	1980	TCTGTCAATTTGCAACTCCCACTTGTATTTTGAACAGGCAATGGAATAAGAAAAATAA	2039
QY	3069	GTAATATTTGTCTCAAG 3084	
Db	2040	GTAATATTTGTCTCAAG 2055	

```

US-09-966-459A-4
; Sequence 4, Application US/09966459A
; Publication No. US20030022237A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, J.N.
; APPLICANT: MINTIER, G.
; APPLICANT: RAMAATHAN, C.S.
; APPLICANT: HAWKEN, D.R.
; APPLICANT: CACACE, A.
; APPLICANT: BARBER, L.
; APPLICANT: KORNAKNER, M.G.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRTM4,
; FILE REFERENCE: D0039NP
; CURRENT APPLICATION NUMBER: US/09/966,459A
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,833
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/261,776
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/305,351
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/313,202
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2034
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-966-459A-4

```

## RESULT 13

Query Match	63.5%	Score 1992.6	DB 11	Length 2034
Best Local Similarity	99.6%	Pred. No. 0		
Matches 2029	Conservative 0	Mismatches 4	Indels 4	Gaps 3
QY	1087	GTGTGAGTATCAAACTCTTTTCATCTCAGAGTCTGTATTCAGATTTTATGTATAC	1144	
Db	1	GTGTGAGTATCAAACTCTTTTCATCTCAGAGTCTGTATTCAGATTTTATGTATAC	60	
QY	1147	ATTTTGGAAAGCAGTATTCAGAAAAAATTTCCTTAATAAAAATACACTCAATCCT	1206	
Db	61	ATTTTGGAAAGCAGTATTCAGAAAAAATTTCCTTAAT-AAAAATACACTCAATCCT	119	
QY	1207	TCAAAATGAAACGTGTGTGGGAAATCTCCATTTTTCATATATATTTCTCTGTGTTT	1266	
Db	120	TCAAAATGAAACGTGTGTGGGAAATCTCCATTTTTCATATATATTTCTCTGTGTTT	179	
QY	1267	CTTGCTACATATTAATTAATACCTCGATAGTGTGTGGAGGTTATTACTTTTC	1322	
Db	180	CTTGCTACATATTAATTAATACCTCGATAGTGTGTGTGGAGGTTATTACTTTTC	239	
QY	1327	ATTTTACATGCAAGTCCAAATCTTAACGTCTTCTACTGATGGTTACAGCAATTCGAAT	1386	
Db	240	ATTTTACATGCAAGTCCAAATCTTAACGTCTTCTACTGATGGTTACAGCAATTCGAAT	299	
QY	1387	AAGATGCTATGATCTAGAGAACTTTGGCCAAAGGCTTAAGACGGGAAAGAAATTAAC	1448	
Db	300	AAGATGCTATGATCTAGAGAACTTTGGCCAAAGGCTTAAGACGGGAAAGAAATTAAC	359	
QY	1447	ACAGATATTAATAAATGAGATATCTAGCTTAAACCTTAACCTTCTTCAGAACTCC	1506	
Db	360	ACAGATATTAATAAATGAGATATCTAGCTTAAACCTTAACCTTCTTCAGAACTCC	419	
QY	1507	CACCACTTGGAGCTCAGAAAAATGCTGCTTCAAAATGAATCTTCTACAGAAATAATA	1566	
Db	420	CACCACTTGGAGCTCAGAAAAATGCTGCTTCAAAATGAATCTTCTACAGAAATAATA	479	
QY	1567	ATTTTCTCTCGACACTTACGACTTAAAGGGAAGATTTGAATAAAGCTTGAATAAGT	1626	
Db	480	ATTTTCTCTCGACACTTACGACTTAAAGGGAAGATTTGAATAAAGCTTGAATAAGT	539	
QY	1627	ACATTTACTACGTTAATGAAAGTTGACACACTGTTCTGAGAGTTTCAACAGCATATGGA	1686	



QY 172 TTGATCTTAATAGGCTCCCTGGTTTGAAGAAGCTCACTTGTGTGGCTTCCCATG 231  
DB 121 TTGATCTTAATAGGCTCCCTGGTTTGAAGAAGCTCACTTGTGTGGCTTCCCATG 180  
QY 232 TGTCCCTTACCTTATGCTGTAGTAACTTGACATCATGATGTTGGGACT 291  
DB 181 TGTCCCTTACCTTATGCTGTAGTAACTTGACATCATGATGTTGGGACT 240  
QY 292 GAGCAAGCTGTAGTAACTTGATATATTTTGTGATGCTTTGAGGATGACATC 351  
DB 241 GAGCAAGCTGTAGTAACTTGATATATTTTGTGATGCTTTGAGGATGACATC 300  
QY 352 CTCATCTCAGCTGATGATGAGCAAAAGTGGGATCTTGTGTTAATTCACATAC 411  
DB 301 CTCATCTCAGCTGATGATGAGCAAAAGTGGGATCTTGTGTTAATTCACATAC 360  
QY 412 ATCCAGTTTGAATGCTGTGTGTAAGATTTTGTGATGATGCTTATCTGGATGAA 471  
DB 361 ATCCAGTTTGAATGCTGTGTGTAAGATTTTGTGATGATGCTTATCTGGATGAA 420  
QY 472 TCCAGAGTGTGCTGCTGCTGCTTGTGATGCTTATGAGGATGCTGATGCTGCTG 531  
DB 421 TCCAGAGTGTGCTGCTGCTGCTTGTGATGCTTATGAGGATGCTGATGCTGCTG 480  
QY 532 CATGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 591  
DB 481 CATGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 592 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651  
DB 541 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
QY 652 AATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 711  
DB 601 AATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660  
QY 712 ATCCGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771  
DB 661 ATCCGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
QY 772 CTTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 831  
DB 721 CTTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 780  
QY 832 GCCCAGGCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891  
DB 781 GCCCAGGCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
QY 892 GTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 951  
DB 841 GTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900  
QY 952 CCCGCTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1011  
DB 901 CCCGCTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
QY 1012 GAGAGGCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1071  
DB 961 GAGAGGCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
QY 1072 GCTTCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1131  
DB 1021 GCTTCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
QY 1132 GATTTTATGTTTACATTTTGAAGAAGATTTTGAAGATTTTCTTAAATTA 1191  
DB 1081 GATTTTATGTTTACATTTTGAAGAAGATTTTGAAGATTTTCTTAAATTA 1139  
QY 1192 TACAACTCAGATCTTCAATATGAAAGTGTGGGATCTCCTTAAATTA 1251  
DB 1140 TACAACTCAGATCTTCAATATGAAAGTGTGGGATCTCCTTAAATTA 1199

QY 1252 TTCTTCTTGTGTTTCTGCTACATATTAATTAATTAATTAATTAATTAATTAAT 1311  
DB 1200 TTCTTCTTGTGTTTCTGCTACATATTAATTAATTAATTAATTAATTAATTAAT 1259  
QY 1312 GGGTATTAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 1354  
DB 1260 GGGTATTAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 1302

RESULT 15  
US-09-780-669-916  
Sequence 916, Application US/09780669  
Patent No. US2002005197A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jianshun  
APPLICANT: Dillon, Devin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqun  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Derrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
APPLICANT: Hurst, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C24  
CURRENT APPLICATION NUMBER: US/09/780, 669  
CURRENT FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 943  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 916  
LENGTH: 1302  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-780-669-916

Query Match 41.1%; Score 1287.8; DB 9; Length 1302;  
Best Local Similarity 99.8%; Pred. No. 3.9e-303;  
Matches 1300; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 52 TGACAAAGGGGCTACACATCTTCCATACGCTTGAAGCTCTACCTGCGCTGGCTGG 111  
DB 1 TGACAAAGGGGCTACACATCTTCCATACGCTTGAAGCTCTACCTGCGCTGGCTGG 60  
QY 112 TCACAGTTCAAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATG 171  
DB 61 TCACAGTTCAAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATG 120  
QY 172 TTCATCTTAATAGGCTCCCTGGTTTGAAGAAGCTGATGCTTGTGGCTTCCCATG 231  
DB 121 TTCATCTTAATAGGCTCCCTGGTTTGAAGAAGCTGATGCTTGTGGCTTCCCATG 180  
QY 232 TGTCCCTTACCTTATGCTGTAGTAACTTGACATCATGATGTTGGGACT 291  
DB 181 TGTCCCTTACCTTATGCTGTAGTAACTTGACATCATGATGTTGGGACT 240  
QY 292 GAGCAAGCTGTAGTAACTTGATATATTTTGTGATGCTTTGAGGATGACATC 351  
DB 241 GAGCAAGCTGTAGTAACTTGATATATTTTGTGATGCTTTGAGGATGACATC 300  
QY 352 CTCATCTCAGCTGATGATGAGCAAAAGTGGGATCTTGTGTTAATTCACATAC 411



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 16:05:18 / Search time 6213 Seconds  
(without alignments)  
12257.638 Million cell updates/sec

Title: US-10-017-066A-1

Sequence: 1 cagagagagctgtatcttcagc.....aaaaaaaaaaaaaaaaaaaaa 3136

Scoring table: IDENTITY NJC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

EST:  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gse\_hum:\*  
18: em\_gse\_inv:\*  
19: em\_gse\_pin:\*  
20: em\_gse\_vrt:\*  
21: em\_gse\_fun:\*  
22: em\_gse\_mam:\*  
23: em\_gse\_mus:\*  
24: em\_gse\_pro:\*  
25: em\_gse\_rnd:\*  
26: em\_gse\_png:\*  
27: em\_gse\_vrt:\*  
28: gb\_gse1:\*  
29: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	701	22.4	702	13	EX092732 EX092732
2	684.8	21.8	753	9	AU138663 AU138663
3	606.8	19.3	765	28	AF101565 AF101565
4	568.6	18.1	664	10	BF679294 BF679294

5	547.8	17.5	779	10	BG541042	BG541042 602570205
6	544.6	17.4	559	9	AF694767	AF694767 we22910.x
7	519.4	15.6	862	10	BF674120	BF674120 602137641
8	498.4	15.0	505	9	AF1871596	AF1871596 we28b05.x
9	471.8	15.0	483	9	AF138131	AF138131 UI-H-B11-
10	464.4	14.8	466	9	AF771398	AF771398 hnt5907.x
11	444.6	14.2	660	10	BB653241	BB653241 BB653241
12	443	14.1	448	9	AA988520	AA988520 cr84a03.s
13	442	14.1	445	10	BE550795	BE550795 7529409.x
14	442	14.1	447	9	AF631017	AF631017 tx4e412.x
15	441	14.1	450	10	BF433082	BF433082 7n2412.x
16	440	14.0	443	9	AF1457247	AF1457247 ck06c12.x
17	432.2	13.8	632	10	BB665936	BB665936 BB665936
18	400.2	12.8	705	9	AF106391	AF106391 um27b11.y
19	399.4	12.7	507	14	CD288243	CD288243 3 lg_abd
20	398.4	12.7	400	13	BX283988	BX283988 BX283988
21	379.4	12.1	454	9	AF157695	AF157695 AU157695
22	378	12.1	386	9	AA531196	AA531196 nj48e04.s
23	355	11.3	1964	11	AK036356	AK036356 Mus muscu
24	355	11.3	3410	11	AK028467	AK028467 Mus muscu
25	343.8	11.0	353	9	AA622418	AA622418 nc45f04.s
26	336.4	10.7	338	10	BG212555	BG212555 RST32147
27	302.4	9.6	583	4	BX525010	BX525010 RZPD Mus
28	284.2	9.1	1086	28	AF101706	AF101706 AF101706
29	281.8	9.0	296	9	AA689245	AA689245 nt90b06.s
30	280	8.9	447	9	AW235635	AW235635 xm20g08.x
31	274.2	8.7	795	29	BZ256668	BZ256668 CH230-334
32	260.4	8.3	662	10	BB550626	BB550626 BB550626
33	254	7.9	462	10	BF002049	BF002049 7g98c09.x
34	246.2	7.9	301	9	AW078525	AW078525 xa89a05.x
35	200.2	6.4	569	10	AW962764	AW962764 B257374837
36	200.4	6.4	773	28	BH112117	BH112117 RPCI-24-2
37	199.4	6.4	695	28	BH043372	BH043372 RPCI-24-3
38	194.8	6.2	820	28	BH046083	BH046083 RPCI-24-2
39	191.4	6.1	705	28	AF101881	AF101881 AF101881
40	187.8	6.0	798	28	BZ612040	BZ612040 WACH19TF
41	187.8	6.0	821	28	BZ195145	BZ195145 CH230-249
42	184.2	5.9	798	28	BH082934	BH082934 RPCI-24-9
43	181.4	5.8	588	28	AZ006959	AZ006959 RPCI-23-3
44	176.6	5.6	859	29	BZ221157	BZ221157 CH230-312
45	176	5.6	784	29	BZ248123	BZ248123 CH230-325

## ALIGNMENTS

RESULT 1  
LOCUS EX092732 702 bp mRNA linear EST 23-JAN-2003  
DEFINITION BX092732 NCI CGAP LUS Homo sapiens cDNA clone IMAGE:1602508, mRNA sequence.

ACCESSION BX092732  
KEYWORDS BX092732.1 GI:27825459

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Ebert L., Heil O., Hennig S., Neubert P., Partsch E., Peters M., Radelof U., Schneider D. and Korn B.

TITLE Human Unigeneset - RZPD3

COMMENT Unpublished

Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD: IMAGP981054063

RZPDLB, I.M.A.G.E. cDNA Clone Collection;

Human Unigeneset - RZPD3 (RZPDLB No. 972)

http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi/response?libNo=972

Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de

This clone is available royalty-free from RZPD.  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
M13r. Primer sequence: TTTCACACAGGAAACAGCTATGAC.  
Location/Qualifiers

## FEATURES

source

```

1..702
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE381054063" ; IMAGE:1602508"
/tissue_type="cardioid"
/lab_host="DH10B"
/clone_id="NCI_CGAP_Lus"
/notes="Organ: lung; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
neuroendocrine lung cardioid, and was then primed with a
Not I - oligo(dT) primer. Double-stranded cDNA was ligated
to Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p773 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      186 a      139 c      142 g      234 t      1 others
ORIGIN

```

Query Match 22.4%; Score 701; DB 13; Length 702;  
Best Local Similarity 99.9%; Pred. No. 6,4e-64;

Matches 701; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

2211 GAAATGACAGCTTTGAGATTAACATTTGACCTTTGAGTGTGCTGTGAGTGAAGTGA 2270
1 GAATGGACAGCTTTGAGATTAACATTTGACCTTTGAGTGTGCTGTGAGTGAAGTGA 60
2271 GGGATCTTCAGACCATGCTTTATTGGGGCTTTGGCAGTATGAAACAGGACCTTGA 2330
61 GGGATCTTCAGACCATGCTTTATTGGGGCTTTGGCAGTATGAAACAGGACCTTGA 120
2331 GACCAGAAAGCAATCTGACTTAGAGGAGGATCAGGCAATTTTGTCTGAGGGGCTA 2390
121 GACCGAGAAAGCAATCTGACTTAGAGGAGGATCAGGCAATTTTGTCTGAGGGGCTA 180
2391 TTACCAAGGTTTAAAGTTTCAATCTTCACAGAGATGACACAGAGTTTAAACAGAAA 2450
181 TTACCAAGGTTTAAAGTTTCAATCTTCACAGAGATGACACAGAGTTTAAACAGAAA 240
2451 CTCAATATTAACAATTAACAATGATGATATATATGATGATGATGATGATGATGATGAT 2510
241 CTCAATATTAACAATTAACAATGATGATATATATGATGATGATGATGATGATGATGAT 300
2511 AATCTCAGGTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2570
301 AATCTCAGGTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
2571 TATCATATTTGGAATGCTATTAATTAATGTTGCTGCTGAGACTGTAACCCATGA 2630
361 TATCATATTTGGAATGCTATTAATTAATGTTGCTGCTGAGACTGTAACCCATGA 420
2631 GGGACAGTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2690
421 GGGACAGTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
2691 TGAATCCCCAGCAAGTGGCTAGAAATATATGTTTATGCTTGAACCGGTTATTTT 2750
481 TGAATCCCCAGCAAGTGGCTAGAAATATATGTTTATGCTTGAACCGGTTATTTT 540
2751 TCAATCAACCTGATTCCTTCTGCTGCTGACATAGCAGGCAATTTTCAAGGCTTCTT 2810
541 TCAATCAACCTGATTCCTTCTGCTGCTGACATAGCAGGCAATTTTCAAGGCTTCTT 600
2811 GAGTTGGGTTATTAATTTCTGGCCCTTCTTCAATGATGATGATGATGATGATGATGAT 2870
601 GAGTTGGGTTATTAATTTCTGGCCCTTCTTCAATGATGATGATGATGATGATGATGAT 660

```

QY 2871 ATTTCTATACCTGCTCATTAACCTCCCATGTCAGGCTT 2912  
Db 661 ATTTCTATACCTGCTCATTAACCTCCCATGTCAGGCTT 702

## RESULT 2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

BASE COUNT

247 a

145 c

122 g

236 t

3 others

Query Match

21.8%; Score 684.8; DB 9; Length 753;

Best Local Similarity 97.5%; Pred. No. 2.9e-62;

Matches 727; Conservative 0; Mismatches 14; Indels 5; Gaps 3;

```

1029 GATTCGACAGGCAATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
3 GATTCGACAGGCAATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 62
1089 GTGAGTGAATCAACTCTTTTCAATTCAGAGTCTGATTCAGATTTTAATGTTAAAT 1148
63 GTGAGTGAATCAACTCTTTTCAATTCAGAGTCTGATTCAGATTTTAATGTTAAAT 122
1149 TTGGAAACAGATTCAGAAATTTTCTTAATTAATTAATTAATTAATTAATTAATTAAT 1208
123 TTGGAAACAGATTCAGAAATTTTCTTAATTAATTAATTAATTAATTAATTAATTAAT 181
1209 AAATATGAACCTGTTGGGAAATCTCAATTTTCAATTAATTAATTTTCTTTGTTTCT 1268
182 AAATATGAACCTGTTGGGAAATCTCAATTTTCAATTAATTAATTTTCTTTGTTTCT 241
1269 TGTACATATTAATTAATTAATTCCTGATGATGATGATGATGATGATGATGATGATGAT 1328
242 TGTACATATTAATTAATTAATTCCTGATGATGATGATGATGATGATGATGATGATGAT 301
1329 TTACCATGACATCAATCAATCTGCTTCACTGATGATGATGATGATGATGATGATGAT 1388

```

FEATURES	LOCATION/Qualifiers
QY	2477 GATGCAATATAGGGGAAGTTGATTCATTTCTTTTGTGTAATCCAGGTTGCCGTATATGAGATT 2536
Db	1 GATGCAATATAGGGGAAGTTGATTCATTTCTTTTGTGTAATCCAGGTTGCCGTATATGAGATT 60
QY	2537 CCTATATACATCTTTCATCCCTTTTGTGAATGATATCATATTTTGAAGAAATGCCATTTTAA 2596
Db	61 CCTATATACATCTTTCATCCCTTTTGTGAATGATATCATATTTTGAAGAAATGCCATTTTAA 120
QY	2597 TACTGTATTTTGGCTGCTGAGACTGTAAAGCCATGAGGGCAGCTGTTATTATTGAATGTGCAT 2656
Db	121 TACTGTATTTTGGCTGCTGAGACTGTAAAGCCATGAGGGCAGCTGTTATTATTGAATGTGCAT 180
QY	2657 CTCTGTTTCATCATTTGACGCTGCTTTGGCTGCATCATATTGAATCCCCCAGAGAAAGTGCCTGAA 2716
Db	181 CTCTGTTTCATCATTTGACGCTGCTTTGGCTGCATCATATTGCCAGAAAGTGCCTGAA 240
QY	2717 CATATATAGTGTATAGCTTGAACACCGGTTATTTTTCATCAAACTGATTCCTTGTCCT 2776
Db	241 CATATATAGTGTATAGCTTGAACACCGGTTATTTTTCATCAAACTGATTCCTTGTCCT 300
QY	2777 GACACATATAGCAGGAGCAATTTTCAGGCTCTTTGATGTTGGGTATTTATTAATTTGCGCC 2836
Db	301 GACACATATAGCAGGAGCAATTTTCAGGCTCTTTGATGTTGGGTATTTATTAATTTGCGCC 360
QY	2837 ATTAATCTTCCATGTGAGTGAAGTGAAGTGAATGTCATTTTATACCTGAGCTCATAAAAACC 2896
Db	361 ATTAATCTTCCATGTGAGTGAAGTGAAGTGAATGTCATTTTATACCTGAGCTCATAAAAACC 420
QY	2897 TCCCATGTGACAGCTTTTCATGTTGACACTTAAATGTGACTTTGGAGAGTATGTGTACACA 2956
Db	421 TCCCATGTGACAGCTTTTCATGTTGACACTTAAATGTGACTTTGGAGAGTATGTGTACACA 480
QY	2957 GAGTAAATTCACACAGAACCTCGATTTTCTGAAAAAACTGTGCAGAGCAAACTCTGTGCAT 3016
Db	481 GAGTAAATTCACACAGAACCTCGATTTTCTGAAAAAACTGTGCAGAGCAAACTCTGTGCAT 540
QY	3017 TTGCAACTCCCACTTGTTATTGTATCGAGAGCAGTTGATTAAGTAAAAATTAAGTACTATT 3076
Db	541 TTGCAACTCCCACTTGTTATTGTATCGAGAGCAGTTGATTAAGTAAAAATTAAGTACTATT 600
QY	3077 GTGTCAAGAAAAAATAA 3094
Db	601 GTGTCAAGTCTCTGAAAA 618
RESULT 4	
LOCUS	Bf679294 664 bp mRNA linear EST 21-DEC-2000
DEFINITION	602133488F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294532 5',
ACCESSION	mRNA sequence.
VERSION	Bf679294
KEYWORDS	Bf679294.1 GI:11953189
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (bases 1 to 664)
COMMENT	NIH-MGC <a href="http://mgi.nci.nih.gov/">http://mgi.nci.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: CLONTECH Laboratories, Inc. cDNA Library Preparation: CLONTECH Laboratories, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: L1CM1143 row: P column: 21 High quality sequence stop: 601.



source

1. .664

/organism="Homo sapiens"

/mol\_type="rRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4294532"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NIH MGC 83"

/note="Organ: Prostate; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctcggc); Site 2: SfiI (ggccattatggc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCCGACG-3' (where B = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 137 a 187 c 147 g 193 t

ORIGIN

Query Match 18.1%; Score 568.6; DB 10; Length 664;  
Best Local Similarity 99.0%; Pred. No. 3.5e-50;  
Matches 593; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

43 GAGGAGAGCTGACAAAGGGGGTCAACATTCCTTCATAGAGCTTCACTGCGC 102  
1 GAGGAGAGCTGACAAAGGGGGTCAACATTCCTTCATAGAGCTTCACTGCGC 60  
103 TGGTCTGTGACATTCAGCTTCTTCAATGATGGATCCCATGGCAATGATCACT 162  
61 TGGTCTGTGACATTCAGCTTCTTCAATGATGGATCCCATGGCAATGATCACT 120  
163 GCTACATCTTCATCTCAATAGAGCTTCTTCAATGATGGATCCCATGGCAATGATCACT 222  
121 GCTACATCTTCATCTCAATAGAGCTTCTTCAATGATGGATCCCATGGCAATGATCACT 180  
223 TTTCCATTTGCTGCTCTTCACTTATGCTGCTAGGATCACTTCACTTCACTT 282  
181 TTTCCATTTGCTGCTCTTCACTTATGCTGCTAGGATCACTTCACTTCACTT 240  
283 GTGGGAGCTGAGCAGAGCTGATGAGGCGCATGATATTTCTTTCAGAGCTTTCAGC 342  
241 GTGGGAGCTGAGCAGAGCTGATGAGGCGCATGATATTTCTTTCAGAGCTTTCAGC 300  
343 ATTGACATCTTCATCTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 402  
301 ATTGACATCTTCATCTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 360  
403 TCCACTACATCCAGTTGATGCTGCTGCTAAGATTTTGGCATCCACTCTTATCT 462  
361 TCCACTACATCCAGTTGATGCTGCTGCTAAGATTTTGGCATCCACTCTTATCT 420  
463 GGCATGGAATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 522  
421 GGCATGGAATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479  
523 CCACGCGCCCATGCGACAGTACTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582  
480 CCACGCGCCCATGCGACAGTACTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539  
583 GTGGTGGCGGGGGGTGACATGATGG-CACCCCTTCTGCTTCACTCAAGAGAGCTGCGCT 640  
540 GTGGTGGCGGGGGGTGACATGATGGCAACCCCTTCTGCTTCACTCAAGAGAGCTGCGCT 598

RESULT 5  
BS541042 779 bp mRNA linear EST 03-APR-2001  
LOCUS 602570205F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4694856 5'  
DEFINITION  
RNA sequence.  
ACCESSION  
BS541042  
VERSION  
BS541042.1 GI:13533275  
KEYWORDS  
EST.

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE  
1 (bases 1 to 779)  
NIH-MGC <http://mgi.nci.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
TITLE  
AUTHORS  
COMMENT  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: CLONTECH Laboratories, Inc.  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
Plate: LCM1519 row: 1 column: 01

FEATURES

source

Location/Qualifiers

1. .779

/organism="Homo sapiens"

/mol\_type="rRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4694856"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NIH MGC 77"

/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctcggc); Site 2: SfiI (ggccattatggc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCCGACGATG-3' (where B = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH MGC library."

BASE COUNT 145 a 205 c 198 g 231 t

ORIGIN

Query Match 17.5%; Score 547.8; DB 10; Length 779;  
Best Local Similarity 91.8%; Pred. No. 4.5e-48;  
Matches 614; Conservative 0; Mismatches 47; Indels 8; Gaps 3;

1 CAGAGAGCTGATTTCACTGAGAGCTGCGAGAGCTTCTTGGAGAGAGCTGAGCAAG 60  
17 CAGAGAGCTGATTTCACTGAGAGCTGCGAGAGCTTCTTGGAGAGAGCTGAGCAAG 76  
61 GGGGTCAACATCTCTTCATACGGTTGAGGCTTACCTGCTGCTGCTGCTGCTGCT 120  
77 GGGGTCAACATCTCTTCATACGGTTGAGGCTTACCTGCTGCTGCTGCTGCTGCT 136  
121 AGCTTCTTCATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
137 AGCTTCTTCATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 196  
181 ATAGAGCTTCTCTGTTTGAAGAGGCTCACTTCTGTTGCTTCCATTGCTGCTC 240  
197 ATAGAGCTTCTCTGTTTGAAGAGGCTCACTTCTGTTGCTTCCATTGCTGCTC 256  
241 TACCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
257 TACCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 316  
301 CTGATGAGGCCAGATATATATTTCTTTCAGAGCTTTCAGAGCTTTCAGAGCTTTC 360  
317 CTGATGAGGCCAGATATATATTTCTTTCAGAGCTTTCAGAGCTTTCAGAGCTTTC 376  
361 ACCTTATTCATGCGCCAAATGCTGCGCATCTTCTGTTCAATTCACTACATCAAGTT 420  
377 ACCTTATTCATGCGCCAAATGCTGCGCATCTTCTGTTCAATTCACTACATCAAGTT 436  
421 GATGCTGTCTGATACAGATTTTTCATCAGTCTTATCTGAGAGATTCACAGTG 480

Db 437 GAGCTGTCTCTGACGATGTTTGGCCATCCCTTATCTGAGATCCACAGG 496  
 Qy 481 CTCTGGCCATGCTTTTGAACCGTATGTGGCCATCTGTACCCAGTGCACAGCA 540  
 Db 497 CTGTGGCCATGCTTTTGAACCG-TATGTGGCCATCTGTACCCAGTGCACAGCA 555  
 Qy 541 GTCCTTA-CGTTGCTGCTGTGCACAAA-----TTGCTGTGCTGTGTGTGCGG 593  
 Db 556 GTACTTACCTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 615  
 Qy 594 GCGTGCATGATGACACCCCTTCTGTCTCATCAAGACGCTCCCTTGTGCGCTCCA 653  
 Db 616 GGTGACCTGTGTGACCCCTTCTGTCTCATCAAGACGCTCCCTTGTGCGCTCCA 675  
 Qy 654 TATCTTTC 662  
 Db 676 TTCTTTCC 684

RESULT 6  
 A1694767 559 bp mRNA linear EST 18-DEC-1999  
 LOCUS A1694767  
 DEFINITION we42910.x1 NCI\_CGAP\_Lu24 Homo sapiens cDNA clone IMAGE:2343810 3',  
 mRNA sequence.  
 ACCESSION A1694767  
 VERSION A1694767.1 GI:4982667  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 COMMENT Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/bbnp/image/image.html  
 Insert Length: 661 Std. Error: 0.00  
 Seq. primer: -40UP from Gibco  
 High quality sequence stop: 417.  
 Location/Qualifiers

FEATURES  
 source

1. 559  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2343810"  
 /tissue\_type="cardioid"  
 /lab\_host="DH10B"  
 /clone\_1ib="NCI CGAP Lu24"  
 /note="Organ: lung; Vector: pUT73D-Pac (Pharmacia) with a  
 modified polylinker; Plasmid DNA from the normalized  
 library NCI CGAP Lu5 was prepared, and ss circles were  
 made in vitro. Following HAP purification, this DNA was  
 used as tracer in a subtractive hybridization reaction.  
 The driver was PCR-amplified cDNAs from a pool of 5,000  
 clones made from the same library (clones1414920-1417991 and 1520904-1522439). Subtraction by Bento  
 Soares and M. Fatima Bonaldo."  
 BASE COUNT 183 a 99 c 110 g 167 t  
 ORIGIN

Query Match 17.4%; Score 544.6; DB 9; Length 559;  
 Best Local Similarity 98.4%; Pred. No. 1.2e-47;

Matches 550; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2543 ACATGCTTCAATCCCTTTGTATAGATATCATATTGGAAAGCCTATTATCTTG 2602  
 Db 559 ACATGCTTCAATCCCTTTGTATAGATATCATATTGGAAAGCCTATTATCTTG 500  
 Qy 2603 TATTTGCTGCTGAATGTAAAGCCATGAGGACCTGTTATTATGATGATCTCTGT 2662  
 Db 499 TATTTGCTGCTGAATGTAAAGCCATGAGGACCTGTTATTATGATGATCTCTGT 440  
 Qy 2663 TCATCATGATGCTGCTTTTGTCTATCATTTGAATCCCCAGAAAGTCCCTGAAATAT 2722  
 Db 439 TCATCATGATGCTGCTTTTGTCTATCATTTGAATCCCCAGAAAGTCCCTGAAATAT 380  
 Qy 2723 AGTCTTATGCTTACACCCGTTATTTTTCATCAACCTGATTCCTGTGCTGACAC 2782  
 Db 379 AGTCTTATGCTTACACCCGTTATTTTTCATCAACCTGATTCCTGTGCTGACAC 320  
 Qy 2783 ATAGCCAGGCAATTTTCCAGCTCTTTGAGTGGTATTTATTAATTCGGCCATTACT 2842  
 Db 319 ATAGCCAGGCAATTTTCCAGCTCTTTGAGTGGTATTTATTAATTTGGCCATTACT 260  
 Qy 2843 TCCATATGATGAGTGAAGTGAATGCAATTTCTATACCTGGCTCAATAAACCTCCCAT 2902  
 Db 259 TCCATATGATGAGTGAAGTGAATGCAATTTCTATACCTGGCTCAATAAACCTCCCAT 200  
 Qy 2903 GTGACGCTTTGATGATGATTAATGTGACTGTGGAGCTATGTGTACAGAGATTA 2962  
 Db 199 GTGACGCTTTGATGATGATTAATGTGACTGTGGAGCTATGTGTACAGAGATTA 140  
 Qy 2963 ATCCACGAGAGCCCTGGAATTTCTGAAAAAATGTGCAAGCCAACTGTCTATTGGCA 3022  
 Db 139 ATCCACGAGAGCCCTGGAATTTCTGAAAAAATGTGCAAGCCAACTGTCTATTGGCA 80  
 Qy 3023 CTCCCACTTGTATTTGTACAGGACAGTGTGATGATGAATAAATAAGTACTATTGTCTCA 3082  
 Db 79 CTCCCACTTGTATTTGTACAGGACAGTGTGATGATGAATAAATAAGTACTATTGTCTCA 20  
 Qy 3083 AGAAAAAATTTTAAAAA 3101  
 Db 19 AAAAAAATTTTAAAAA 1

RESULT 7  
 BF674130 862 bp mRNA linear EST 21-DEC-2000  
 LOCUS BF674130  
 DEFINITION 60213764.F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4273634 5',  
 mRNA sequence.  
 ACCESSION BF674130  
 VERSION BF674130.1 GI:11948025  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 TITLE NIH-MGC http://mgi.nci.nih.gov/  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: CLONETECH Laboratories, Inc.  
 CDNA Library Preparation: CLONETECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L10M1090 row: b column: 11  
 High quality sequence stop: 628.  
 Location/Qualifiers

FEATURES  
 source  
 1. 862  
 /organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4273834"
/lab_host="DH10B (T1 phage-resistant)"
/clone_id="NTH_MGC_83"
/notes="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccgctcggcc); Site_2: SfiI (ggccatcagcc)
; 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT      185 a      233 c      223 t
ORIGIN

```

```

Query Match      16 6%; Score 519.4; DB 10; Length 862;
Best Local Similarity 97.1%; Pred. No. 3.8e-45;
Matches 603; Conservative 0; Mismatches 11; Indels 7; Gaps 7;

```

```

QY 41 TGGAGAAAGACTGACAAAGGAGGAGTACACATTCCTCCATACGCTTACGCTTACCTG 100
DB 2 TGGAGAAAGACTGACAAAGGAGGAGTACACATTCCTCCATACGCTTACGCTTACCTG 61
QY 101 CCTGATGCTGTCACAGTTCAGCTTCTTCATGATGATGATGCCAATGGCAATGATCCA 160
DB 62 CCTGATGCTGTCACAGTTCAGCTTCTTCATGATGATGATGCCAATGGCAATGATCCA 121
QY 161 GTGTACATACATCTTCAATTAAGAGGCTCCTGCTTTAGAGAGGCTCAATTCGCTTGG 220
DB 122 GTGTACATACATCTTCAATTAAGAGGCTCCTGCTTTAGAGAGGCTCAATTCGCTTGG 181
QY 221 CCTTCCCATGCTGCTCCTCTTACCTTATGCTGTGCTAGTAACTTACATCACTTACA 280
DB 182 CCTTCCCATGCTGCTCCTCTTACCTTATGCTGTGCTAGTAACTTACATCACTTACA 241
QY 281 TTGTGCGAGTGAAGACAGGCTGATGAGGAGGATGATATTTCTTGTGATGCTTTCAG 340
DB 242 TTGTGCGAGTGAAGACAGGCTGATGAGGAGGATGATATTTCTTGTGATGCTTTCAG 301
QY 341 GCATGATGATCCTTCACTTCCACCTCATCATGCCCAATGCTGGCATCTTGTGTTCA 400
DB 302 GCATGATGATCCTTCACTTCCACCTCATCATGCCCAATGCTGGCATCTTGTGTTCA 361
QY 401 ATTCCACATACATCCAGTT-TGATGCT-TGCTGCTTACAGAT-TTTTGCCATCCACTCT 457
DB 362 ATTCCACATACATCCAGTT-TGATGCT-TGCTGCTTACAGAT-TTTTGCCATCCACTCT 421
QY 458 TATCTGAGATGATTCACAGTGTGCTGGCCATGGC-TTTTGACCGCTATGTGGCCATC 516
DB 422 TATCTGAGATGATTCACAGTGTGCTGGCCATGGC-TTTTGACCGCTATGTGGCCATC 481
QY 517 TGTACCCACATGCGGCATGCGACAGTATTAAGTGTGCTGTGT-CACCAAAATGGTGT 575
DB 482 TGTACCCACATGCGGCATGCGACAGTATTAAGTGTGCTGTGT-CACCAAAATGGTGT 541
QY 576 GGTGCTGTGTGTGGGGGGGCTGACATGATGACACCCCTTCTCTTCAACAGAGCT 635
DB 542 GGTGCTGTGTGTGGGGGGGCTGACATGATGACACCCCTTCTCTTCAACAGAGCT 600
QY 636 GCCCTTTCGCGGCTTCATAT 656
DB 601 G-CCTTTCGCGGCTTCATATCT 620

```

```

RESULT 8
A1871596/c 505 bp mRNA linear EST 18-DEC-1999
LOCUS A1871596
DEFINITION mRNA sequence.
ACCESSION A1871596

```

```

VERSION A1871596.1 GI:5545645
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE 1
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

```

```

COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
www.bio.liml.gov/bdrip/image/image.html
Insert Length: 938 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 458.

```

```

FEATURES
source
1..505
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2342385"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_id="NCI-CGAP Lu24"
/notes="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI CGAP Lu5 was prepared, and ss circles were
used in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (clones
1414920-1417951 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT      165 a      95 c      109 g      136 t
ORIGIN

```

```

Query Match      15 9%; Score 498.4; DB 9; Length 505;
Best Local Similarity 99.8%; Pred. No. 7.7e-43;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2585 ATGCTATTATATCTTATTTGCTGCTGACATGATGAGGCACTGTTATTT 2644
DB 505 ATGCTATTATATCTTATTTGCTGCTGACATGATGAGGCACTGTTATTT 446
QY 2645 ATTGAATGTCATCTCTGTTATCATATGATGCTCTTGTCTCATATGAAATCCCAACA 2704
DB 445 ATTGAATGTCATCTCTGTTATCATATGATGCTCTTGTCTCATATGAAATCCCAACA 386
QY 2705 AAGTGCCTAGAACATATAGTCTTATGCTTACACACCGGTTATTTTCATCAAACTGAT 2764
DB 385 AAGTGCCTAGAACATATAGTCTTATGCTTACACACCGGTTATTTTCATCAAACTGAT 326
QY 2765 TCTTCTGTCTGGAACACATAGCCAGCAATTTTCCACCTTCTTTGAGTTGGGATTAAT 2824
DB 325 TCTTCTGTCTGGAACACATAGCCAGCAATTTTCCACCTTCTTTGAGTTGGGATTAAT 266
QY 2825 TAAATTCGACCATTAATTCATGATGATGAGTGAAGTGCATGTCATTTCTATACCTGG 2884
DB 265 TAAATTCGACCATTAATTCATGATGATGAGTGAAGTGCATTTCTATACCTGG 206
QY 2885 CTCATTAACCTCCCATGTCAGGCTTTTCATGTTGACATTAATGTACTTGGAACT 2944
DB 205 CTCATTAACCTCCCATGTCAGGCTTTTCATGTTGACATTAATGTACTTGGAACT 146

```

QY 2945 ATGTGTTACAGAGTAATATACCAAGACCTGATTTCTGAAAAAATCTGTGACAGACCA 3004  
 DB 145 ATGTGTTACAGAGTAATATACCAAGACCTGATTTCTGAAAAAATCTGTGACAGACCA 86  
 QY 3005 AACCTGTCTATTTGCAACTCCCACTTGTATTTGTACAGAGAGTTGGATTAATGAAAAA 3064  
 DB 85 AACCTGTCTATTTGCAACTCCCACTTGTATTTGTACAGAGAGTTGGATTAATGAAAAA 26  
 QY 3065 TAAAGTACTATTTGTCTAAG 3084  
 DB 25 TAAAGTACTATTTGTCTAAG 6

RESULT 9  
 AM138131/c 483 bp mRNA linear EST 29-OCT-1999  
 LOCUS  
 DEFINITION UI-H-B11-acy-a-08-0-UI.s1 NCI\_CGAP\_Sub3 Homo sapiens cDNA clone  
 IMAGE:2715926 3', mRNA sequence.  
 ACCESSION AM138131  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 483)  
 NCI\_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL  
 COMMENT  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 The sequence contained an oligo-dt track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. cDNA library preparation: W.B. Soares Lab Clone distribution:  
 NCI\_CGAP clone distribution information can be found through the  
 I.M.A.G.E. Consortium/ILM: at:  
 www.bio.ilm.gov/bdrrp/image/image.html  
 Seq primer: M13 Forward  
 POLY(A)=yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..483  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2715926"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_1lb="NCI\_CGAP\_Sub3"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; The  
 NCI\_CGAP Sub3 library is a subtracted library derived from  
 the NCI CGAP Sub1 library, which is a subtracted library  
 derived from B1. B1 constitutes a mixture of 21  
 normalized or subtracted NCI CGAP libraries: NCI\_CGAP\_Co4  
 , NCI\_CGAP\_Pr22, NCI\_CGAP\_Pr28, NCI\_CGAP\_Co10,  
 NCI\_CGAP\_Co16, NCI\_CGAP\_Kid5, NCI\_CGAP\_Kid12,  
 NCI\_CGAP\_Kid3, NCI\_CGAP\_Co8, NCI\_CGAP\_Co11, NCI\_CGAP\_Kid12,  
 NCI\_CGAP\_Br2, NCI\_CGAP\_Co8, NCI\_CGAP\_Co11, NCI\_CGAP\_Let12,  
 NCI\_CGAP\_Br23, NCI\_CGAP\_Lus, NCI\_CGAP\_Lu24,  
 NCI\_CGAP\_Lu19, NCI\_CGAP\_GC4, NCI\_CGAP\_GC6,  
 NCI\_CGAP\_Br25. These 21 libraries were pooled and a  
 single-stranded DNA preparation of the resulting mixture  
 was used as a tracer in a subtractive hybridization with  
 a driver whose composition is detailed below:  
 NCI\_CGAP\_Kid3 pool 1 LLM 3334-3337, 3682-3683,  
 3798-3803 (IMAGE Clonids 132376-132391, 1456008-1456775  
 , 1500552-1502855); NCI\_CGAP\_Kid5 pool 1 LLM 3338-3342  
 , 3722-3725, 3776-3778 (IMAGE Clonids 133912-132831,  
 1471368-1472803, 1492104-149255); NCI\_CGAP\_Lus pool 1  
 LLM 3575-3582, 3851-3854 (IMAGE Clonids 1414920-1417991,  
 1520904-1522439); NCI\_CGAP\_GC4 pool 1 LLM 3164-3167,  
 3716-3720, 3733-3735 (IMAGE Clonids 1257096-1258631,

BASE COUNT 149 a 89 c 102 g 143 t  
 ORIGIN  
 Query Match 15.0%; Score 471.8; DB 9; Length 483;  
 Best local similarity 98.6%; Pred. No. 4.5e-40;  
 Matches 476; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2625 CCATGAGGCACTGTTTATATGATGTCATCTCTGTCATCATTTGCTCTTTGCT 2684  
 DB 483 CCATGAGGCACTGTTTATATGATGTCATCTCTGTCATCATTTGCTCTTTGCT 424  
 QY 2685 CATCATTAATCCGCCAGCAAGGCTCAGAAATATAGGCTTATGCTTGAACCGGT 2744  
 DB 423 CATCATTAATCCGCCAGCAAGGCTCAGAAATATAGGCTTATGCTTGAACCGGT 364  
 QY 2745 TATTTTTCATCAAACTGATTCCTCTGCTGGAACATATGACGAGCAATTTTCAGCC 2804  
 DB 363 TATTTTTCATCAAACTGATTCCTCTGCTGGAACATATGACGAGCAATTTTCAGCC 304  
 QY 2805 TTCTTTGAGTGGGATATTAATTTCTGGGATTAATTTCTTCAATGATGAGTGAAGTACA 2864  
 DB 303 TTCTTTGAGTGGGATATTAATTTCTGGGATTAATTTCTTCAATGATGAGTGAAGTACA 244  
 QY 2865 TGTGCAATTTCTAATACCTGCTCATTAACCCCTCCAGTGCAGCCTTTCAATGTCAT 2924  
 DB 243 TGTGCAATTTCTAATACCTGCTCATTAACCCCTCCAGTGCAGCCTTTCAATGTCAT 184  
 QY 2925 TAAATGTACCTGGGAAAGCTATGTGTACACAGATTAATACAGAAAGCTGATTTCT 2984  
 DB 183 TAAATGTACCTGGGAAAGCTATGTGTACACAGATTAATACAGAAAGCTGATTTCT 124  
 QY 2985 GAAAAAATCTGTGACAGCAAACTGTCATTTGCAACTCCACTGATATTTGTACAG 3044  
 DB 123 GAAAAAATCTGTGACAGCAAACTGTCATTTGCAACTCCACTGATATTTGTACAG 64  
 QY 3045 GCAAGTGGATTAAGTAAATTAAGTACTATTTGTCAAGAAAAA 3104  
 DB 63 GCAAGTGGATTAAGTAAATTAAGTACTATTTGTCAAGAAAAA 4  
 QY 3105 AAA 3107  
 DB 3 AAA 1

RESULT 10  
 AM771398/c 466 bp mRNA linear EST 04-MAY-2000  
 LOCUS  
 DEFINITION hms5907.x1 NCI\_CGAP\_Co17 Homo sapiens cDNA clone IMAGE:3027612 3',  
 mRNA sequence.  
 ACCESSION AM771398  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 466)  
 NCI\_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL  
 Unpublished





```

VERSION      BE550795.1  GI:9792487
KEYWORDS     EST
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE        1 (bases 1 to 445)
JOURNAL      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
              Unpublished
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
              Emmert-Buck, M.D., Ph.D.
              CDNA Library Preparation: M. Bento Soares, Ph.D.
              CDNA Sequencing by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/HLNLI, send email to:
              info@image.llnl.gov
              Seq primer: -400P from Gibco.
              Location/Qualifiers
                1..445
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:3232504"
                /tissue_type="carcinoid"
                /lab_host="DH10B"
                /clone_id="NCI CGAP Lu24"
                /note="Organ: lung; Vector: pUT3D-Pac (Pharmacia) with a
                modified polylinker; Plasmid DNA from the normalized
                library NCI CGAP Lys was prepared, and ss circles were
                used as tracer in a subtractive hybridization reaction.
                The driver was PCR-amplified cDNAs from a pool of 5,000
                clones made from the same library (cloneids
                141920-1417991 and 1520904-1522439). Subtraction by Bento
                Soares and M. Fatima Ronaldo."
BASE COUNT   144 a      82 c      97 g      122 t
ORIGIN
Query Match 14.1%; Score 442; DB 10; Length 445;
Best Local Similarity 100.0%; Pred. No. 5.8e-37;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2643 TTATTGAATGTCATCTGTCATCATCTGCTCTTTGCTCATCTGAATCCCCCAG 2702
Db 445 TTATTGAATGTCATCTGTCATCATCTGCTCTTTGCTCATCTGAATCCCCCAG 386
Cy 2703 CAAAGTGCCTAGAACATATATAGTCTTATGCTTACACCGGTTATTTTCATCAAACTG 2762
Db 385 CAAAGTGCCTAGAACATATATAGTCTTATGCTTACACCGGTTATTTTCATCAAACTG 326
Cy 2763 ATTGCTTCTGTCGACACATAGCCAGGCAATTTCCACCTTCTTGAGTTGGGTATT 2822
Db 325 ATTGCTTCTGTCGACACATAGCCAGGCAATTTCCACCTTCTTGAGTTGGGTATT 266
Cy 2823 ATTAAATCTGCGCATTAATCTTCCATGTAGTGGAATGACATGCAATTTCTATACCT 2882
Db 265 ATTAAATCTGCGCATTAATCTTCCATGTAGTGGAATGACATGCAATTTCTATACCT 206
Cy 2883 GGCTCATAAACCTTCCCATGTGACGCTTTCATGTTGACATTAATGTGACTTGGGAAG 2942
Db 205 GGCTCATAAACCTTCCCATGTGACGCTTTCATGTTGACATTAATGTGACTTGGGAAG 146
Cy 2943 CTATGTGTTACACAGAGTAATACAGAGAGCCGATTTCTGAAAAAAGTGGCAGAG 3002
Db 145 CTATGTGTTACACAGAGTAATACAGAGAGCCGATTTCTGAAAAAAGTGGCAGAG 86
Cy 3003 CAAAGCTCTGTGATTTGCAACTCCCACTTGTATTGTACAGGACAGTTGATTAAGTAA 3062

```

```

Db 85 CAAAGCTCTGTGATTTGCAACTCCCACTTGTATTGTACAGGACAGTTGATTAAGTAA 26
Cy 3063 AATTAAGTACTATTTGTCAG 3084
Db 25 AATTAAGTACTATTTGTCAG 4

RESULT 14
AI631017/c
LOCUS
DEFINITION tx54el2.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2273422 3',
mRNA sequence.
ACCESSION AI631017
VERSION AI631017.1 GI:4682347
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE 1 (bases 1 to 447)
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
              Unpublished
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
              Emmert-Buck, M.D., Ph.D.
              CDNA Library Preparation: M. Bento Soares, Ph.D.
              CDNA Sequencing by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/HLNLI at:
              www.bio.llnl.gov/bbtp/image/image.html
              Insert Length: 539 Std Error: 0.00
              Seq primer: -400P from Gibco.
              Location/Qualifiers
                1..447
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:2273422"
                /tissue_type="carcinoid"
                /lab_host="DH10B"
                /clone_id="NCI CGAP Lu24"
                /note="Organ: lung; Vector: pUT3D-Pac (Pharmacia) with a
                modified polylinker; Plasmid DNA from the normalized
                library NCI CGAP Lys was prepared, and ss circles were
                used as tracer in a subtractive hybridization reaction.
                The driver was PCR-amplified cDNAs from a pool of 5,000
                clones made from the same library (cloneids
                141920-1417991 and 1520904-1522439). Subtraction by Bento
                Soares and M. Fatima Ronaldo."
BASE COUNT   145 a      82 c      97 g      123 t
ORIGIN
Query Match 14.1%; Score 442; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.8e-37;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2643 TTATTGAATGTCATCTGTCATCATCTGCTCTTTGCTCATCTGAATCCCCCAG 2702
Db 447 TTATTGAATGTCATCTGTCATCATCTGCTCTTTGCTCATCTGAATCCCCCAG 388
Cy 2703 CAAAGTGCCTAGAACATATATAGTCTTATGCTTACACCGGTTATTTTCATCAAACTG 2762
Db 387 CAAAGTGCCTAGAACATATATAGTCTTATGCTTACACCGGTTATTTTCATCAAACTG 328
Cy 2763 ATTGCTTCTGTCGACACATAGCCAGGCAATTTCCACCTTCTTGAGTTGGGTATT 2822
Db 327 ATTGCTTCTGTCGACACATAGCCAGGCAATTTCCACCTTCTTGAGTTGGGTATT 268

```

QY 2823 ATTAATCTGGCCATTCTTCATGAGTGAAGTGAATGTCATTTCTTACT 2882  
 DB 267 ATTAATCTGGCCATTCTTCATGAGTGAAGTGAATGTCATTTCTTACT 208  
 QY 2883 GGGTCATTAACCCCTCCATGTCAGCCCTTCATGTCATTAATGTCATGGGAAG 2942  
 DB 207 GGGTCATTAACCCCTCCATGTCAGCCCTTCATGTCATTAATGTCATGGGAAG 148  
 QY 2943 CTATGTGTACACAGTAATATCAGCAGACCTGGATTTCTGAAAAAAGTCGACAGC 3002  
 DB 147 CTATGTGTACACAGTAATATCAGCAGACCTGGATTTCTGAAAAAAGTCGACAGC 88  
 QY 3003 CAACCTCTGTCATTTGCACTCCCACTTGATTTGTCAGAGGAGTGTAGTAA 3062  
 DB 87 CAACCTCTGTCATTTGCACTCCCACTTGATTTGTCAGAGGAGTGTAGTAA 28  
 QY 3063 AATAAGTACTATTGTGTCAAG 3084  
 DB 27 AATAAGTACTATTGTGTCAAG 6

## RESULT 15

LOCUS BF433082 450 bp mRNA linear EST 29-NOV-2000  
 DEFINITION 7n24f12.x1 NCI\_CGAP\_Lu24 Homo sapiens CDNA clone IMAGE:3565606 3',  
 mRNA sequence.  
 ACCESSION BF433082  
 VERSION BF433082.1 GI:11445245  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS 1 (bases 1 to 450)  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapdb-remail.nih.gov  
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL, send email to:  
 info@image.lnl.gov

Seq primer: -40UP from Glibco.

FEATURES  
 Source Location/Qualifiers  
 1..450

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3565606"  
 /tissue\_type="carcinoid"  
 /lab\_host="DH10B"  
 /clone\_id="NCI\_CGAP\_Lu24"  
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a  
 modified polylinker; Plasmid DNA from the normalized  
 library NCI CGAP Lu5 was prepared, and ss circles were  
 made in vitro. Following HAP purification, this DNA was  
 used as tracer in a subtractive hybridization reaction.  
 The driver was PCR-amplified cDNAs from a pool of 5,000  
 clones made from the same library (cloneids  
 141920-1417991 and 1520904-1522439). Subtraction by Bento  
 Soares and M. Fatima Bernaldo."

BASE COUNT 144 a 83 c 98 g 125 t

Query Match 14.1%; Score 441; DB 10; Length 450;  
 Best Local Similarity 100.0%; Freq. No. 7.3e-37;  
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2644 TATTGATGTCATCTCTGTTCATGTCAGTGTCTTGGCTATGATGATCCCCAGC 2703  
 DB 450 TATTGATGTCATCTCTGTTCATGTCAGTGTCTTGGCTATGATGATCCCCAGC 391  
 QY 2704 AAGTGCCTAGAACTAATAGTCTTANGCTTGAACCCGGTATTTTCATCAAACTGA 2763  
 DB 390 AAGTGCCTAGAACTAATAGTCTTANGCTTGAACCCGGTATTTTCATCAAACTGA 331  
 QY 2764 TTCCTCTGCTGGAACACATAGCCAGGCAATTTTCAAGCCCTCTTGGAGTGGATTA 2823  
 DB 330 TTCCTCTGCTGGAACACATAGCCAGGCAATTTTCAAGCCCTCTTGGAGTGGATTA 271  
 QY 2824 TTAATTCGGCCATTACTCTTCATGTCAGTGAAGTGAATGTCATTAACCTG 2883  
 DB 270 TTAATTCGGCCATTACTCTTCATGTCAGTGAAGTGAATGTCATTAACCTG 211  
 QY 2884 GGTCAATAAACCCCTCCATGTCAGCCCTTCATGTCATTAATGTCATTTGGGAAGC 2943  
 DB 210 GGTCAATAAACCCCTCCATGTCAGCCCTTCATGTCATTAATGTCATTTGGGAAGC 151  
 QY 2944 TATGTGTACACAGTAATATCAGCAGAACCTGGATTTCTGAAAAAAGTCGACAGCC 3003  
 DB 150 TATGTGTACACAGTAATATCAGCAGAACCTGGATTTCTGAAAAAAGTCGACAGCC 91  
 QY 3004 AAACCTCTGTCACTTTGCACTCCCACTTGATTTGTGACAGGCACTTGATTAAGTGA 3063  
 DB 90 AAACCTCTGTCACTTTGCACTCCCACTTGATTTGTGACAGGCACTTGATTAAGTGA 31  
 QY 3064 AATAAGTACTATTGTGTCAAG 3084  
 DB 30 AATAAGTACTATTGTGTCAAG 10

Search completed: February 9, 2004, 21:17:14  
 Job time : 6223 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 16:05:17 ; Search time 1154 Seconds  
(without alignments)  
11501.936 Million cell updates/sec

Title: US-10-017-066a-1  
Perfect score: 3136  
Sequence: 1 cagagagagctgtatttcagt.....aaaaaaaaaaaaaaaaaaaaa 3136

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

GenBank: 1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_in:\*  
18: em\_mu:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_pr:\*  
26: em\_ro:\*  
27: em\_scs:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_tam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3136	100.0	3136	6 AX107037	AX107037 Sequence
2	3136	100.0	3136	6 AX549148	AX549148 Sequence
3	3028.8	96.5	3077	9 HSM804438	AL833127 Homo sapi
4	2969.6	94.7	175073	2 AC105936	AC105936 Homo sapi
5	2969.6	94.7	179177	9 AC090719	AC090719 Homo sapi
6	2028.4	64.7	2061	6 AX201144	AX201144 Sequence
7	2028.4	64.7	2061	6 AX267943	AX267943 Sequence
8	2028.4	64.7	2061	6 BD160111	BD160111 Primer fo
9	2028.4	64.7	2061	9 AK023643	AK023643 Homo sapi
10	1940.8	61.9	179664	2 AC025249	AC025249 Homo sapi
11	1287.8	41.1	1302	6 AX201143	AX201143 Sequence
12	1287.8	41.1	1302	6 AX267942	AX267942 Sequence
13	1186.4	37.8	1225	9 BC022401	BC022401 Homo sapi
14	1186.2	37.8	1354	6 AB064553	AB064553 Sequence
15	1186.2	37.8	1354	9 AB065787	AB065787 Homo sapi
16	966.4	30.8	968	6 AX338395	AX338395 Sequence
17	964.2	30.7	968	6 AX338397	AX338397 Sequence
18	962.2	30.7	968	6 AX338399	AX338399 Sequence
19	955.4	30.5	957	6 AX128501	AX128501 Sequence
20	955.4	30.5	957	6 AX201145	AX201145 Sequence
21	955.4	30.5	957	6 AX267944	AX267944 Sequence
22	955.4	30.5	957	6 AX448799	AX448799 Sequence
23	955.4	30.5	957	6 BD144354	BD144354 Novel G-P
24	952.4	30.4	954	6 AX201146	AX201146 Sequence
25	952.4	30.4	954	6 AX241660	AX241660 Sequence
26	952.4	30.4	954	6 AX267945	AX267945 Sequence
27	884.2	28.2	339098	2 AC096030	AC096030 Rattus no
28	766.8	24.5	954	10 AY072993	AY072993 Mus muscu
29	684.8	21.8	753	6 BD149982	BD149982 Primer fo
30	404	12.9	427	6 AX107041	AX107041 Sequence
31	381.2	12.2	399	6 BD076875	BD076875 5' EST OF
32	379.4	12.1	945	6 BD154949	BD154949 Primer fo
33	379.4	12.1	945	6 AX241763	AX241763 Sequence
34	379.4	12.1	963	6 AX448797	AX448797 Sequence
35	379.4	12.1	963	6 BD144351	BD144351 Novel G-P
36	379.4	12.1	1024	6 AX551415	AX551415 Sequence
37	379.4	12.1	175073	2 AC105936	AC105936 Homo sapi
38	379.4	12.1	179177	9 AC090719	AC090719 Homo sapi
39	377.4	12.0	963	6 AR261055	AR261055 Sequence
40	377.4	12.0	963	6 AR278586	AR278586 Sequence
41	377.4	12.0	963	6 AX141036	AX141036 Sequence
42	377.4	12.0	963	6 AX200896	AX200896 Sequence
43	377.4	12.0	963	6 AX267552	AX267552 Sequence
44	377.4	12.0	963	6 AX549376	AX549376 Sequence
45	377.4	12.0	1015	6 AX556222	AX556222 Sequence

## ALIGNMENTS

RESULT 1	AX107037	3136 bp	DNA	linear	PAT 30-APR-2001
LOCUS	AX107037	Sequence 1 from Patent WO0125434.			
DEFINITION	AX107037				
ACCESSION	AX107037.1	GI:13922563			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE					
AUTHORS	Raitano, A.B., Afar, D.E., Jakobovits, A., Farris, M., Hubert, R.S.,				
	Mitchell, S.C. and Saffran, D.C.				
TITLE	G protein-coupled receptor up-regulated in prostate cancer and uses				



QY 1801 TGGGGTCATACAGATATAAAATTTAAAAAAGAGCTTATGCCAATCTCATAGA 1860  
 DB 1801 TGGGGTCATACAGATATAAAATTTAAAAAAGAGCTTATGCCAATCTCATAGA 1860  
 QY 1861 TGTGGAGAACTGTTCAGAGACCAACAGGGTATGAGGTTAGAGATTTCCAGAGCTTAC 1920  
 DB 1861 TGTGGAGAACTGTTCAGAGACCAACAGGGTATGAGGTTAGAGATTTCCAGAGCTTAC 1920  
 QY 1921 ATTTTCTAGAGAGATTTATTTCTTCTCATCATCCAGTGTGATTTAGAGATTTTC 1980  
 DB 1921 ATTTTCTAGAGAGATTTATTTCTTCTCATCATCCAGTGTGATTTAGAGATTTTC 1980  
 QY 1981 CTGGCAACAGAACTCATGGCTTTATCCCATGATTTGCTTATGCTGCTCAAT 2040  
 DB 1981 CTGGCAACAGAACTCATGGCTTTATCCCATGATTTGCTTATGCTGCTCAAT 2040  
 QY 2041 GCCAATTACCTGTCTGTGAGAGAGATTTCTAGTTCACCATTAATGGAATTTCTTA 2100  
 DB 2041 GCCAATTACCTGTCTGTGAGAGAGATTTCTAGTTCACCATTAATGGAATTTCTTA 2100  
 QY 2101 TTGAGAAAGTCTGATAGGCTTTATAGCAAGTTATTTTAAAAAGTTCCATAGTGA 2160  
 DB 2101 TTGAGAAAGTCTGATAGGCTTTATAGCAAGTTATTTTAAAAAGTTCCATAGTGA 2160  
 QY 2161 TTCTGATAGGCACTGAGTTAGAGACCAACAGTTATGATGGAAAGTATGAGATGCAAG 2220  
 DB 2161 TTCTGATAGGCACTGAGTTAGAGACCAACAGTTATGATGGAAAGTATGAGATGCAAG 2220  
 QY 2221 TCTTGAAGATTAACATTTGCTTTGAGTGTGATCTGATCTGAAAGTATGAGGAACTTC 2280  
 DB 2221 TCTTGAAGATTAACATTTGCTTTGAGTGTGATCTGATCTGAAAGTATGAGGAACTTC 2280  
 QY 2281 AGGACCATGCTTTATTTGGGGCTTTGTGCAATGAGAACAGGACTTTGAGACCAAGAAA 2340  
 DB 2281 AGGACCATGCTTTATTTGGGGCTTTGTGCAATGAGAACAGGACTTTGAGACCAAGAAA 2340  
 QY 2341 GCATCTGATAGGCACTGAGTTAGAGACCAACAGTTATGATGGAAAGTATGAGATGCAAG 2400  
 DB 2341 GCATCTGATAGGCACTGAGTTAGAGACCAACAGTTATGATGGAAAGTATGAGATGCAAG 2400  
 QY 2401 TTATATGCTTATCTCTTCAACAGATATGACAAACAGTGTAAACAAAGAACTCAATATAC 2460  
 DB 2401 TTATATGCTTATCTCTTCAACAGATATGACAAACAGTGTAAACAAAGAACTCAATATAC 2460  
 QY 2461 AAATACATAAAACATGATCATATATGAGTATGATCTTTCTTTTCAATCTCTCAGG 2520  
 DB 2461 AAATACATAAAACATGATCATATATGAGTATGATCTTTCTTTTCAATCTCTCAGG 2520  
 QY 2521 TTCCCTGATATGATCTCTTCAACAGATATGATCTTTCTTTTCAATCTCTCAGG 2580  
 DB 2521 TTCCCTGATATGATCTCTTCAACAGATATGATCTTTCTTTTCAATCTCTCAGG 2580  
 QY 2581 GGAAGGCTATTTATCTGTATTTGCTGTGACCTGTAAGCCCAAGAGGCACTGTT 2640  
 DB 2581 GGAAGGCTATTTATCTGTATTTGCTGTGACCTGTAAGCCCAAGAGGCACTGTT 2640  
 QY 2641 TATTATGATGATCATCTCTGTTCATCATGATCTCTCTTGTATGATGATCAATCAATCC 2700  
 DB 2641 TATTATGATGATCATCTCTGTTCATCATGATCTCTCTTGTATGATGATCAATCAATCC 2700  
 QY 2701 AGGAAAGTGTGATCATATATGATGCTTATGCTTGAACAGGCTTATTTTCAATCAAC 2760  
 DB 2701 AGGAAAGTGTGATCATATATGATGCTTATGCTTGAACAGGCTTATTTTCAATCAAC 2760  
 QY 2761 TGAATCTCTGTCTGTGACCAATATGAGCCAGGCAATTTTCCAGCTCTTTGATGAGGTA 2820  
 DB 2761 TGAATCTCTGTCTGTGACCAATATGAGCCAGGCAATTTTCCAGCTCTTTGATGAGGTA 2820  
 QY 2821 TTAATTAATTTGGGCACTTACTTCAATGTGATGAGAGTGAATGCAATTTCTATAC 2880  
 DB 2821 TTAATTAATTTGGGCACTTACTTCAATGTGATGAGAGTGAATGCAATTTCTATAC 2880  
 QY 2881 CTGGCTCATATAAACCTCCATGTGAGGCTTTCAATGTTGACATTAATGAGACTTGGGA 2940

DB 2881 CTGGCTCATATAAACCTCCATGTGACAGCTTTTCAATTAATGATGATGATGAG 2940  
 QY 2941 AGCTATGTGTTACACAGATTAATACACAGAGCTGATTTCTGAAAAAAGCTGTGACA 3000  
 DB 2941 AGCTATGTGTTACACAGATTAATACACAGAGCTGATTTCTGAAAAAAGCTGTGACA 3000  
 QY 3001 GCCAACCCTGTGATTTGCAACCTCCACTGATTTGTAGAGGAGCTTGGATTAATGTA 3060  
 DB 3001 GCCAACCCTGTGATTTGCAACCTCCACTGATTTGTAGAGGAGCTTGGATTAATGTA 3060  
 QY 3061 AAAATTAAGTACTATTTGTGCAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 3120  
 DB 3061 AAAATTAAGTACTATTTGTGCAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 3120  
 QY 3121 AAAAAAAAAAAAAAAAAA 3136  
 DB 3121 AAAAAAAAAAAAAAAAAA 3136  
 RESULT 2  
 AX549148  
 LOCUS AX549148 3136 bp DNA linear PAT 26-NOV-2002  
 DEFINITION Sequence 433 from Patent WO02061087.  
 ACCESSION AX549148  
 VERSION AX549148.1 GI:25813884  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 Butler, G.C., Roush, C.L. and Brown, J.P.  
 Antigenic peptides, such as for G protein-coupled receptors  
 (GPCRs), antibodies thereto, and systems for identifying such  
 antigenic peptides  
 Patent: WO 02061087-A 433 08-AUG-2002;  
 Lifespan Biosciences, Inc. (US)  
 FEATRES  
 source  
 1. 3136  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 864 a 674 c 622 g 976 t  
 ORIGIN  
 Query Match 100.0%; Score 3136; DB 6; Length 3136;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAGAGAGGCTGTATTTCAAGTGCAGCTGCGCAACCTTTTGGAGAGAAAGCTGACAAAG 60  
 DB 1 CAGAGAGGCTGTATTTCAAGTGCAGCTGCGCAACCTTTTGGAGAGAAAGCTGACAAAG 60  
 QY 61 GGGGTCAACATTCCTTCCATACAGTTCAGCTCTTCTGAGGCTGCTGCTGCTGCTGCTG 120  
 DB 61 GGGGTCAACATTCCTTCCATACAGTTCAGCTCTTCTGAGGCTGCTGCTGCTGCTGCTG 120  
 QY 121 AGCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
 DB 121 AGCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
 QY 181 ATAGGCTCCCTGCTGTTAGAGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 DB 181 ATAGGCTCCCTGCTGTTAGAGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 241 TACCTTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 DB 241 TACCTTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 301 CTGCATAGAGCCCATGATATATTTCTTTCATGCTTTCAGGCAATGACATCCCATCTCC 360  
 DB 301 CTGCATAGAGCCCATGATATATTTCTTTCATGCTTTCAGGCAATGACATCCCATCTCC 360





[illegible]

Db	1919	TTGTTGATTTTGGAAATTTCTGGCAACGAACTATGGCTTTAATCCCACTAGCTATTGC	1978
Oy	2022	TTATGTCTGGTCCAAATTGCAATTAACCTGTGTCTTGGAGAAATGATTTCTAGTTCA	2081
Db	1979	TTATGTCTGGTCCAAATTGCAATTAACCTGTGTCTTGGAGAAATGATTTCTAGTTCA	2038
Oy	2082	CCATTATGGAAAGATTTCTTATTCAGAAAGTCTGCATAGGGCTTATAGCAAGTATTTATT	2144
Db	2039	CCATTATGGAAAGATTTCTTATTCAGAAAGTCTGCATAGGGCTTATAGCAAGTATTTATT	2096
Oy	2142	TTAAAGTTCATAGTGTATCTGATAGGACGATGAGTTAGGAGCCACCACTTATGATG	2201
Db	2099	TTAAAGTTCATAGTGTATCTGATAGGACGATGAGTTAGGAGCCACCACTTATGATG	2156
Oy	2202	GGAGGTATGGAATGGCAGGTCTTGAAGTAAATTTGGCTTTTGAAGTGTACTGTAAGCT	2266
Db	2159	GGAGGTATGGAATGGCAGGTCTTGAAGTAAATTTGGCTTTTGAAGTGTACTGTAAGCT	2218
Oy	2262	GGAAAGTGAAGGAATCTTACAGGACCATAGCTTTA TTGGGGGCTTTGTGCAATGGAACAG	2324
Db	2219	GGAAAGTGAAGGAATCTTACAGGACCATAGCTTTA TTGGGGGCTTTGTGCAATGGAACAG	2276
Oy	2322	GGACTTTGAGCCAGGAAAGCAATCTGACTTAGGCATGGGAATAGGCAATTTTGTCTCT	2381
Db	2279	GGACTTTGAGCCAGGAAAGCAATCTGACTTAGGCATGGGAATAGGCAATTTTGTCTCT	2336
Oy	2382	GAGGGGCTATTACCAAGGGTTAATAGGTTTCATTTTCAACGAAATAGCAACAGTGTTA	2444
Db	2339	GAGGGGCTATTACCAAGGGTTAATAGGTTTCATTTTCAACGAAATAGCAACAGTGTTA	2396
Oy	2442	ACCAAGAACTCAAATTTACAAATTAATAAAATGTGATCATATAATGTGTAGTTTCATT	2501
Db	2399	ACCAAGAACTCAAATTTACAAATTAATAAAATGTGATCATATAATGTGTAGTTTCATT	2458
Oy	2502	TTCTTTTTCATCTCTCAGGTTCCTGTATATAGATTTCTCTATTAACATGCTTATATCCCTTT	2566
Db	2459	TTCTTTTTCATCTCTCAGGTTCCTGTATATAGATTTCTCTATTAACATGCTTATATCCCTTT	2518
Oy	2562	TGTATATGATATCATATTTTGGAAATGCTATTTTATCTTGATTTCTGTCTGTGACGTGTA	2621
Db	2519	TGTATATGATATCATATTTTGGAAATGCTATTTTATCTTGATTTCTGTGTGACGTGTA	2578
Oy	2622	AGCCCATGAGGGCACGTGTTTATTAATGAAATGTCACTCTGTTCATCATGACCTGCTTTT	2681
Db	2579	AGCCCATGAGGGCACGTGTTTATTAATGAAATGTCACTCTGTTCATCATGACCTGCTTTT	2638
Oy	2682	GCTCATATGTAATCCCCAGCAAAAGTCCATGAAACATAATAGGCTTATCTTGACACC	2744
Db	2639	GCTCATATGTAATCCCCAGCAAAAGTCCATGAAACATAATAGGCTTATCTTGACACC	2696
Oy	2742	GGTATATTTTTCATCAAACTGTATCTTGTCTGTCTGTAACATATGCAAGCAATTTTCCA	2801
Db	2699	GGTATATTTTTCATCAAACTGTATCTTGTCTGTCTGTAACATATGCAAGCAATTTTCCA	2758
Oy	2802	GCCTCTTTAGTGGGATTTATTTAAATTCGCGCATTACTTCCAAATGTAGTGGAAAGTG	2861
Db	2759	GCCTCTTTAGTGGGATTTATTTAAATTCGCGCATTACTTCCAAATGTAGTGGAAAGTG	2818
Oy	2862	ACATGTGCAATTTCTATACCTGTGCTCAATAAACCTTCCCATGTGCAAGCTTTCAATGTGA	2921
Db	2819	ACATGTGCAATTTCTATACCTGTGCTCAATAAACCTTCCCATGTGCAAGCTTTCAATGTGA	2878
Oy	2922	CATTAAATGTGACCTTGGGAAGCTATGTGTATACACAGAGTAAATATACAGAAAGCTGGATTT	2981
Db	2879	CATTAAATGTGACCTTGGGAAGCTATGTGTATACACAGAGTAAATATACAGAAAGCTGGATTT	2938
Oy	2982	TCTGAAAAAACTGTGACAGACCAAACCTCTGTGATTTGCAACTCCCACTTGTATTTGTAC	3041
Db	2939	TCTGAAAAAACTGTGACAGACCAAACCTCTGTGATTTGCAACTCCCACTTGTATTTGTAC	2998
Oy	3042	GAGGCACTTGAATAGTGAATAATTAAGTCTATTTGTCTCAAGAAAAAATTTTTTTTTTTT	3101





```
Db 174253 CTGCTTCATCAAGAGCTGCCCTTCGCGCTCCAAATCTCTTCCATTCCTACCTGCC 174194
QY 677 TACACCAAGATGTGATGAAGCTGCGCTGTGATGATATCCGGGTCAATGCTGTATGCGC 736
Db 174193 TACACCAAGATGTGATGAAGCTGCGCTGTGATGATATCCGGGTCAATGCTGTATGCGC 174134
QY 737 TATCGTCATCATCTCCGCTATTTGGCTGGATGATCATCTCTCATCTCCCTCTCATATCTGC 796
Db 174133 TATCGTCATCATCTCCGCTATTTGGCTGGATGATCATCTCTCATCTCCCTCTCATATCTGC 174074
QY 797 TATCTTAAAGCTGTGTGGCTTTGACACGTGAAGCCCAAGCCCAAGGCAATTTGCACTT 856
Db 174073 TATCTTAAAGCTGTGTGGCTTTGACACGTGAAGCCCAAGCCCAAGGCAATTTGCACTT 174014
QY 857 GGGCTCTCATGNGTGTGTGTCTGTCATATCTATGATGATACCTTTCACTTGTGTCATGG 916
Db 174013 GGGCTCTCATGNGTGTGTGTCTGTCATATCTATGATGATACCTTTCACTTGTGTCATGG 173954
QY 917 TGCATCGCTTTAGCAAGCGCGTGACTCTCGCTGCGCTCCGTCATCTTGGCCAAATATCTATC 976
Db 173953 TGCATCGCTTTAGCAAGCGCGTGACTCTCGCTGCGCTCCGTCATCTTGGCCAAATATCTATC 173894
QY 977 TGCCTGCTCTCTCTGTGCTCAACCAATTTGCTATGAGTGAAGCAAAAGAGATTGAC 1036
Db 173893 TGCCTGCTCTCTCTGTGCTCAACCAATTTGCTATGAGTGAAGCAAAAGAGATTGAC 173834
QY 1037 AGCGATCCTTGAATTTTCCATGTCGACACAGCGCTTGAGCCCTAGGTCAGTGA 1096
Db 173833 AGCGATCCTTGAATTTTCCATGTCGACACAGCGCTTGAGCCCTAGGTCAGTGA 173774
QY 1097 TCAAACTTCTTTCCATTCAGAGTCTCTGATTCAGATTTTAAATGTTAACTTTTGGAG 1156
Db 173773 TCAAACTTCTTTCCATTCAGAGTCTCTGATTCAGATTTTAAATGTTAACTTTTGGAG 173714
QY 1157 ACAGATTCAGAAAAAATTTTCTTATATAAAAAATCACTGATTCCTTCAATATGA 1216
Db 173713 ACAGATTCAGAAAAAATTTTCTTATATAAAAAATCACTGATTCCTTCAATATGA 173655
QY 1217 AACTGTTGGGGAATCTCCATTTTTCATATTAATTTTCTTCTTGTGTTCTTCTGAT 1276
Db 173654 AACTGTTGGGGAATCTCCATTTTTCATATTAATTTTCTTCTTGTGTTCTTCTGAT 173595
QY 1277 ATAAATTAATAATACCTTACAGTGTGTGTGAGAGGTTATTAATTTTCAATTTTACAT 1336
Db 173594 ATAAATTAATAATACCTTACAGTGTGTGTGAGAGGTTATTAATTTTCAATTTTACAT 173535
QY 1337 GCAGTCCAAATCTAACTGCTTCTACTGATGATGTTTACAGCATTCGTGAGATTAAGATGTA 1396
Db 173534 GCAGTCCAAATCTAACTGCTTCTACTGATGATGTTTACAGCATTCGTGAGATTAAGATGTA 173475
QY 1397 CATCTAGGAACATTTTGCAGAAAGGCTTAAGCAAGGCAAAAGAAATTAACAAGATATA 1456
Db 173474 CATCTAGGAACATTTTGCAGAAAGGCTTAAGCAAGGCAAAAGAAATTAACAAGATATA 173415
QY 1457 ATAAATAGAGTAATCTAGTCTTAAACATATACTTCCCTTCAAGACTCCCAACCAAT 1516
Db 173414 ATAAATAGAGTAATCTAGTCTTAAACATATACTTCCCTTCAAGACTCCCAACCAAT 173355
QY 1517 GGAATCTCAAAAAATGCTGTCTTCAAAATGAATTTTCAAGAGAAATTAATTTTCTC 1576
Db 173354 GGAATCTCAAAAAATGCTGTCTTCAAAATGAATTTTCAAGAGAAATTAATTTTCTC 173295
QY 1577 TGACACATGAGCACTTAAAGGAGAGATGGAAGTAAAGCTTGAATAAGATACATTAACCT 1636
Db 173294 TGACACATGAGCACTTAAAGGAGAGATGGAAGTAAAGCTTGAATAAGATACATTAACCT 173235
QY 1637 ACCTTAATGAAGTGAACAATGCTGTGAGAGTTTTCACAGCATATGACCCCTGTTTT 1696
Db 173234 ACCTTAATGAAGTGAACAATGCTGTGAGAGTTTTCACAGCATATGACCCCTGTTTT 173175
QY 1697 CCTATTAATTTTCTTATCAACCTTAAATTAAGCAAGATATTAATAGACCTCATTTG 1756
Db 173174 CCTATTAATTTTCTTATCAACCTTAAATTAAGCAAGATATTAATAGACCTCATTTG 173115
QY 1757 TACCATGGGAAATATGATTTCAAGTGGGATCAGTGAATTAATTAAGGCTCATACAGTA 1816
Db 173114 TACCATGGGAAATATGATTTCAAGTGGGATCAGTGAATTAATTAAGGCTCATACAGTA 173055
QY 1817 TAAAAATTAAAAAAAGAACTTATGCCAATCTCATGATGTGGAAGAACTGTTA 1876
Db 173054 TAAAAATTAAAAAAAGAACTTATGCCAATCTCATGATGTGGAAGAACTGTTA 172997
QY 1877 GAGAGACCAACAGGGTAGTGGGTAGAGATTTCCAGAGCTTACATTTTCTAGAGAGGT 1936
Db 172996 GAGAGACCAACAGGGTAGTGGGTAGAGATTTCCAGAGCTTACATTTTCTAGAGAGGT 172937
QY 1937 ATTTAATTTCTTCTCATCTCATCAGTGTGTGATTTAGAAATTTCTGGAACAACCTCA 1996
Db 172936 ATTTAATTTCTTCTCATCTCATCAGTGTGTGATTTAGAAATTTCTGGAACAACCTCA 172877
QY 1997 TGGCTTTAATCCCAAGCTAATGGTATTTGCTGATGCTGGTCCCAATTTGCAATTAACCTGTTC 2056
Db 172876 TGGCTTTAATCCCAAGCTAATGGTATTTGCTGATGCTGGTCCCAATTTGCAATTAACCTGTTC 172817
QY 2057 TTGAAAGAGTATTTCTAGTTCACCATTTATGAAAGATTTCTTATTCAGAAAGTCTGAT 2116
Db 172816 TTGAAAGAGTATTTCTAGTTCACCATTTATGAAAGATTTCTTATTCAGAAAGTCTGAT 172757
QY 2117 AGGCTTTATGAGAGTTATTTATTTTAAAGTTCCATAGGTGATTTCTGATAGGCACTA 2176
Db 172756 AGGCTTTATGAGAGTTATTTATTTTAAAGTTCCATAGGTGATTTCTGATAGGCACTA 172697
QY 2177 GGTTAGGAGCCACAGTTATGATGGAAGTATGGAATGGCAGTCTTGAAGATTAACAT 2236
Db 172696 GGTTAGGAGCCACAGTTATGATGGAAGTATGGAATGGCAGTCTTGAAGATTAACAT 172637
QY 2237 GGCCTTTAGAGTGAATCTGTGCTGGAAGAGTGAAGGAATCTTCAAGACATGCTTAT 2296
Db 172636 GGCCTTTAGAGTGAATCTGTGCTGGAAGAGTGAAGGAATCTTCAAGACATGCTTAT 172577
QY 2297 TGGGCTTTGACAGTATGGAACAGGACTTTGAGACAGGAAACCAATCTAGTAAAGC 2356
Db 172576 TGGGCTTTGACAGTATGGAACAGGACTTTGAGACAGGAAACCAATCTAGTAAAGC 172517
QY 2357 ATGGGATCAGGCAATTTTGTCTTCTGAGGGCTATTAACAAAGGTTAATAGGTTTCACT 2416
Db 172516 ATGGGATCAGGCAATTTTGTCTTCTGAGGGCTATTAACAAAGGTTAATAGGTTTCACT 172457
QY 2417 TCAACAGATATGACAACAGTGTAAACCAAGAACTCAAAATTAACAATTAACATGT 2476
Db 172456 TCAACAGATATGACAACAGTGTAAACCAAGAACTCAAAATTAACAATTAACATGT 172397
QY 2477 GATCAATATGTGTAGTTCATTTTCTTCAATCTCAGGTTCCCTGATATGAT 2536
Db 172396 GATCAATATGTGTAGTTCATTTTCTTCAATCTCAGGTTCCCTGATATGAT 172337
QY 2537 CCTATTAACATGCTTCATCCCTTTTGAATGGAATATTTTGAATGCTATTTTA 2596
Db 172336 CCTATTAACATGCTTCATCCCTTTTGAATGGAATATTTTGAATGCTATTTTA 172277
QY 2597 TACTGTATTTGCTGCTGAGCTGAAGCCATGAAGGCACTGTTATTAATGAATGAT 2656
Db 172276 TACTGTATTTGCTGCTGAGCTGAAGCCATGAAGGCACTGTTATTAATGAATGAT 172217
QY 2657 CTCTGTTCATCTTACCTGCTCTTGTCTTCATCATTTGAATTCCTGCTGCT 2716
Db 172216 CTCTGTTCATCTTACCTGCTCTTGTCTTCATCATTTGAATTCCTGCTGCT 172157
QY 2717 CATATATAGTCTTATGCTTGAACCGGTTATTTTCAACACCTGATTCCTGCTGCT 2776
Db 172156 CATATATAGTCTTATGCTTGAACCGGTTATTTTCAACACCTGATTCCTGCTGCT 172097
QY 2777 GAACCATAGCCAGGCAATTTTCCAGCTTCTTGAAGTGGGATTAATTAATTTGCGC 2836
Db 172096 GAACCATAGCCAGGCAATTTTCCAGCTTCTTGAAGTGGGATTAATTAATTTGCGC 172037
```



QY 2837 ATTACTTCAGTGTAGTGGAGTGCATGTGCATTTCTATACCTGGCTCATTAACCC 2896  
 Db 172036 ATTACTTCAGTGTAGTGGAGTGCATGTGCATTTCTATACCTGGCTCATTAACCC 171977  
 QY 2897 TCCCATGTGAGCCTTCATGTGACATTAAATGTGACTGGAGACTTATGTCTTCAACA 2956  
 Db 171976 TCCCATGTGAGCCTTCATGTGACATTAAATGTGACTGGAGACTTATGTCTTCAACA 171917  
 QY 2957 GAGTAAATCAACGAGACCTGTGATTTCTGAAAAAACTGTGCAGAGCCAACTCTGTCAAT 3016  
 Db 171916 GAGTAAATCAACGAGACCTGTGATTTCTGAAAAAACTGTGCAGAGCCAACTCTGTCAAT 171857  
 QY 3017 TTGGAACTCCCACTGTATTTGTATCGAGGAGCTTGATTAAGTAAATTAAGTACTATT 3076  
 Db 171856 TTGGAACTCCCACTGTATTTGTATCGAGGAGCTTGATTAAGTAAATTAAGTACTATT 171797  
 QY 3077 GTGTCAAGAAAAAAA 3094  
 Db 171796 GTGTCAAGTCTGTGAAA 171779

RESULT 5  
 AC090719/c 179177 bp DNA linear PRI 01-FEB-2002  
 LOCUS Homo sapiens chromosome 11, clone RP11-648B16, complete sequence.  
 DEFINITION AC090719  
 AC090719.8 GI:18129429  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 1 (bases 1 to 179177)  
 Homo sapiens chromosome 11, clone RP11-648B16  
 Unpublished  
 2 (bases 1 to 179177)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
 Barna, N., Bastien, V., Boguski, L., Bouckgeater, B., Brown, A.,  
 Camarata, J., Campione, A., Choquet, Y., Colangelo, M., Collins, S.,  
 Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,  
 Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,  
 Gardy, N., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hago, B., Heath, A., Horton, L., Hulme, W., Iller, I., Johnson, R.,  
 Jones, C., Karatas, A., Lacroque, K., Lamazeres, R., Landers, T.,  
 Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,  
 Margis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,  
 McPherson, R., Meldrum, J., Meneus, L., Minova, T., Miura, V.,  
 Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,  
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
 Phunkhang, P., Plerre, N., Pollara, V., Raymond, C., Retter, R.,  
 Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M.,  
 Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P.,  
 Senguez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Sudamanian, A., Talamas, J., Testaye, S., Theodore, J.,  
 Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A.,  
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,  
 Zembek, L., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Submitted (09-MAR-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 179177)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguski, L., Bouckgeater, B.,  
 Brown, A., Camarata, J., Campione, A., Chang, J., Chazaro, B.,  
 Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
 Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,  
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardy, N.,  
 Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hago, B., Horton, L., Hulme, W., Iller, I., Johnson, R., Jones, C.,  
 Karatas, A., Karatas, A., Kells, C., Lacroque, K., Lamazeres, R.,  
 Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,

TITLE  
 JOURNAL  
 COMMENT  
 Submitted (01-FEB-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jan 11, 2002 this sequence version replaced gi:17402791.  
 All repeats were identified using RepeatMasker:  
 Smith, A.P.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

FEATURES  
 source  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seg.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L1849  
 Center clone name: 648\_B\_16

Location/Qualifiers  
 1..179177  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11"  
 /clone="RP11-648B16"  
 /clone\_id="RP11-648B16 Human Male BAC"  
 588..592  
 /note="30 qual SNGC region."  
 complement(814..955)  
 /rpt\_family="L2"  
 1379..1962  
 /rpt\_family="L1MA4"  
 1966..2216  
 /rpt\_family="L1MB1"  
 complement(2229..2541)  
 /rpt\_family="Aluub"  
 2844..2863  
 /rpt\_family="CCAN"  
 complement(2889..3128)  
 /rpt\_family="AluSc"  
 complement(3420..3531)  
 /rpt\_family="MIR"  
 5633..5698  
 /rpt\_family="MIR"  
 5843..5909  
 /rpt\_family="(TTA)n"  
 6121..6215  
 /rpt\_family="L2"  
 6276..6316  
 /rpt\_family="AT-rich"  
 complement(6715..6788)  
 /rpt\_family="MER5A"  
 6843..7000  
 /rpt\_family="L2"  
 complement(7311..7632)  
 /rpt\_family="MSTC"  
 complement(7967..8124)  
 /rpt\_family="L1ME"  
 9014..9090  
 /rpt\_family="MIR"  
 complement(9151..9564)  
 /rpt\_family="MTL1D"



QY 1037 AGGCGATCCTTGAACCTTTCCATGATGGCCACACAGCTTACAGCCCTAGGTGTCAGTGA 1096  
DB 55402 AGGCGATCCTTGAACCTTTCCATGATGGCCACACAGCTTACAGCCCTAGGTGTCAGTGA 55343  
QY 1097 TCAACCTCTTTCCATGAGTCCCTGATTCAGATTTTAACTTAACTTTTGAAG 1156  
DB 55342 TCAACCTCTTTCCATGAGTCCCTGATTCAGATTTTAACTTAACTTTTGAAG 55283  
QY 1157 ACAATATTCAGAAAAAAATTTTCTTAAATAAATAACAATCAGATCCCTCAATATGA 1216  
DB 55282 ACAATATTCAGAAAAAAATTTTCTTAAATAAATAACAATCAGATCCCTCAATATGA 55224  
QY 1217 AACGTGTTGGGGAATCTTCATTTTTCATATTTTCTTCTTGTGTTTCTTGTCAAT 1276  
DB 55223 AACGTGTTGGGGAATCTTCATTTTTCATATTTTCTTCTTGTGTTTCTTGTCAAT 55164  
QY 1277 ATAATTTATTAACCTGATGATGTTGTTGAGGTTATTAATCTTTTCAATTTACAT 1336  
DB 55163 ATAATTTATTAACCTGATGATGTTGTTGAGGTTATTAATCTTTTCAATTTACAT 55104  
QY 1337 GCACTCCAAATCTTAACCTGTTCTACTGATGTTTACAGATTTCTGATTAAGATGTA 1396  
DB 55103 GCACTCCAAATCTTAACCTGTTCTACTGATGTTTACAGATTTCTGATTAAGATGTA 55044  
QY 1397 CATCTAGAGAAATTTGCGCAAGGCTTAAGCAAGGCAAGAAATTAACAGATATTA 1456  
DB 55043 CATCTAGAGAAATTTGCGCAAGGCTTAAGCAAGGCAAGAAATTAACAGATATTA 54984  
QY 1457 AATAAATGAGATTAATCTAGCTTAAATACTAATCTTCTCTTCAAGATCCCAACCAT 1516  
DB 54983 AATAAATGAGATTAATCTAGCTTAAATACTAATCTTCTCTTCAAGATCCCAACCAT 54924  
QY 1517 GGATCTCAGAAAAATGCTGCTTCAAAATGATCTTACAGAAAGAAATTTTCTCTC 1576  
DB 54923 GGATCTCAGAAAAATGCTGCTTCAAAATGATCTTACAGAAAGAAATTTTCTCTC 54864  
QY 1577 TGGACATAGACATTTAAGGGAAGATGGAAGTAAAGCTTGAAGAGATCAATTAAT 1636  
DB 54863 TGGACATAGACATTTAAGGGAAGATGGAAGTAAAGCTTGAAGAGATCAATTAAT 54804  
QY 1637 ACCTTAATGAAGTTGACACATGTTCTGAGATTTTACAGACATTAAGACCTGTTT 1696  
DB 54803 ACCTTAATGAAGTTGACACATGTTCTGAGATTTTACAGACATTAAGACCTGTTT 54744  
QY 1697 CCTATTAAATTTTCTTATCAACCTTTAATTAAGCAAAAGATTAATTAAGACCTGAT 1756  
DB 54743 CCTATTAAATTTTCTTATCAACCTTTAATTAAGCAAAAGATTAATTAAGACCTGAT 54684  
QY 1757 TAGCCATGGGAAATGATGTTGATGAGGATCAGTGAATTAATGAGGTCATACAGTA 1816  
DB 54683 TAGCCATGGGAAATGATGTTGATGAGGATCAGTGAATTAATGAGGTCATACAGTA 54624  
QY 1817 TAAATAATTAATAAATAAATAAAGCTTCAATGCGCAATCTCATATGATGGAAGAACTGTA 1876  
DB 54623 TAAATAATTAATAAATAAATAAAGCTTCAATGCGCAATCTCATATGATGGAAGAACTGTA 54566  
QY 1877 GAGAGACCAACAGGATGAGGTTAGAGATTTCCAGAGTCTTACATTTTCTAGAGAGGT 1936  
DB 54565 GAGAGACCAACAGGATGAGGTTAGAGATTTCCAGAGTCTTACATTTTCTAGAGAGGT 54506  
QY 1937 ATTAAATTTCTTCTCATCTCATCCAGTGTGTAATTTAGAGAAATTTCTGCGCAAGAACTCA 1996  
DB 54505 ATTAAATTTCTTCTCATCTCATCCAGTGTGTAATTTAGAGAAATTTCTGCGCAAGAACTCA 54446  
QY 1997 TGGCTTTAATCCCACTAGCTATTGCTTATGCTGCTGCTCAATTTGCAATTTACCTGTGC 2056  
DB 54445 TGGCTTTAATCCCACTAGCTATTGCTTATGCTGCTGCTCAATTTGCAATTTACCTGTGC 54386  
QY 2057 TTGGAAGAAGATTTTCTAGGTTCACTTAATGGAAGATTTCTTAATTCAGAAATCTGCAT 2116  
DB 54385 TTGGAAGAAGATTTTCTAGGTTCACTTAATGGAAGATTTCTTAATTCAGAAATCTGCAT 54326

QY 2117 AGGCTTAATAGCAATTAATTTTAAAGTTCCATAGGTGATTCGATAGGCAAGTGA 2176  
DB 54325 AGGCTTAATAGCAATTAATTTTAAAGTTCCATAGGTGATTCGATAGGCAAGTGA 54266  
QY 2177 GGTAGGAGGACCAAGATTAATGAGGAAGATGAATGCGAGGCTTGAAGATTAATCT 2236  
DB 54265 GGTAGGAGGACCAAGATTAATGAGGAAGATGAATGCGAGGCTTGAAGATTAATCT 54206  
QY 2237 GGCCTTTTGAATGATCTGATGAGTGAAGATGAGGGAATCTTCAAGCAATGCTTAT 2296  
DB 54205 GGCCTTTTGAATGATCTGATGAGTGAAGATGAGGGAATCTTCAAGCAATGCTTAT 54146  
QY 2297 TGGGCTTTTGTGATGATGAGGACAGGACCTTTGAGCAAGAAAGCAATCTGATAGGC 2356  
DB 54145 TGGGCTTTTGTGATGATGAGGACAGGACCTTTGAGCAAGAAAGCAATCTGATAGGC 54086  
QY 2357 ATGGGAATCAGGCAATTTTGTCTTGAAGGAGCTTATCAAGGCTTAAATAGTTTCAAT 2416  
DB 54085 ATGGGAATCAGGCAATTTTGTCTTGAAGGAGCTTATCAAGGCTTAAATAGTTTCAAT 54026  
QY 2417 TCAACAGATTAATGACACAGTGTTAACCAAGAAATCTCAATTAATTAATTAATCATGT 2476  
DB 54025 TCAACAGATTAATGACACAGTGTTAACCAAGAAATCTCAATTAATTAATTAATCATGT 53966  
QY 2477 GATCATATATGATGATGATGATTTTCTTTTCAATCCCTGAGTCCCTGATATGAT 2536  
DB 53965 GATCATATATGATGATGATGATTTTCTTTTCAATCCCTGAGTCCCTGATATGAT 53906  
QY 2537 CCTATACATGCTTTCATCCCTTTTGTATGATGATATCATATTTGGAATGCTATTTAA 2596  
DB 53905 CCTATACATGCTTTCATCCCTTTTGTATGATGATATCATATTTGGAATGCTATTTAA 53846  
QY 2597 TACTTGAATTTGCTGCTGAGATGATTAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAT 2656  
DB 53845 TACTTGAATTTGCTGCTGAGATGATTAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAT 53786  
QY 2657 CTCTGTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2716  
DB 53785 CTCTGTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53726  
QY 2717 CATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2776  
DB 53725 CATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53666  
QY 2777 GAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2836  
DB 53665 GAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53606  
QY 2837 ATTAATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2896  
DB 53605 ATTAATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53546  
QY 2897 TCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2956  
DB 53545 TCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53486  
QY 2957 GAGTAAATCAACAGAGGCTGATTTCTGAAGAAATCTGTGAGAGCAACCTGTGAT 3016  
DB 53485 GAGTAAATCAACAGAGGCTGATTTCTGAAGAAATCTGTGAGAGCAACCTGTGAT 53426  
QY 3017 TTGCAATCTCCCACTTTGATTTTGTAGAGGCAAGGCTTGAATTAATTAATTAATTAAT 3076  
DB 53425 TTGCAATCTCCCACTTTGATTTTGTAGAGGCAAGGCTTGAATTAATTAATTAATTAAT 53366  
QY 3077 GTGTCAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3094  
DB 53365 GTGTCAAGCTCTGAAAAA 53348

RESULT 6  
AZ201144 AX201144 2061 bp DNA linear PAT 29-AUG-2001  
LOCUS  
DEFINITION Sequence 774 from Patent WO0151633.

ACCESSION AX201144  
VERSION AX201144.1 GI:15390908  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,  
Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W.,  
Stolk, J.A., Skeiky, Y.A., Wang, A. and Mesgher, M.J.  
Compositions and methods for the therapy and diagnosis of prostate  
cancer  
Journal Patent: WO 0151633-A 774 19-JUL-2001,  
CORIXA CORPORATION (US)  
FEATURES  
source location/Qualifiers  
1..2061  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 612 a 383 c 399 g 667 t  
ORIGIN  
Query Match 64.7%; Score 2028.4; DB 6; Length 2061;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;  
QY 1029 GATTGAGACGGGATCCCTTCGATCTTTCATGATGAGCCACACAGCCTTCAGAGCCCTTGGT 1088  
DB 3 GATTGAGACGGGATCCCTTCGATCTTTCATGATGAGCCACACAGCCTTCAGAGCCCTTGGT 62  
QY 1089 GTGAGTGAACAACCTCTTTCATTCAGAGTCCCTGATTCAGATTATATGTTAAT 1148  
DB 63 GTGAGTGAACAACCTCTTTCATTCAGAGTCCCTGATTCAGATTATATGTTAAT 122  
QY 1149 TTTGGAACACGATATTCAGAAAAAATTTCTTATATATATATATATATATATATATATATAT 1208  
DB 123 TTTGGAACACGATATTCAGAAAAAATTTCTTATATATATATATATATATATATATATATAT 181  
QY 1209 AATATGAACCTGTTGGGGAATCTCCATTTTTCATATATATATATATATATATATATATATAT 1268  
DB 182 AATATGAACCTGTTGGGGAATCTCCATTTTTCATATATATATATATATATATATATATATATAT 241  
QY 1289 TGTCTCAT 1328  
DB 242 TGTCTCAT 301  
QY 1329 TTTACCATGACATCCAAATCTAAATGCTTCTTACATGATGTTTACAGCATTCAGAGTAA 1388  
DB 302 TTTACCATGACATCCAAATCTAAATGCTTCTTACATGATGTTTACAGCATTCAGAGTAA 361  
QY 1389 GAATGATCATCTAGAGAACATTTGCCAAGGCTTAAGCAGGCAAGGCAAGGCAAGGCAAGGCAAG 1448  
DB 362 GAATGATCATCTAGAGAACATTTGCCAAGGCTTAAGCAGGCAAGGCAAGGCAAGGCAAGGCAAG 421  
QY 1449 AGAAT 1508  
DB 422 AGAAT 481  
QY 1509 ACCACATGATCTCAGAAAAATGCTGCTTCAAAATGACTTCTTACAGAGAGAAATATAT 1568  
DB 482 ACCACATGATCTCAGAAAAATGCTGCTTCAAAATGACTTCTTACAGAGAGAAATATAT 541  
QY 1569 TTTTCTCTGAGACCTAGCACTTAAGGGGAGATTTGAAGGAGGCTTGAAGGAGGCTTGAAGGAGGCT 1628  
DB 542 TTTTCTCTGAGACCTAGCACTTAAGGGGAGATTTGAAGGAGGCTTGAAGGAGGCTTGAAGGAGGCT 601  
QY 1629 ATTATACCTAGCTTAATGAAGGATGACACATCTTCTGAGAGTTTTCACAGCATATGAGAC 1688  
DB 602 ATTATACCTAGCTTAATGAAGGATGACACATCTTCTGAGAGTTTTCACAGCATATGAGAC 661  
QY 1689 CTGTTTTCTTAT 1748

DB 662 CTGTTTTCTTAT 721  
QY 1749 CCTCATTTAGGACATGGGAAAAATGATGTTTCAGTGGGGATCGATGATTAATGGGGTCA 1808  
DB 722 CCTCATTTAGGACATGGGAAAAATGATGTTTCAGTGGGGATCGATGATTAATGGGGTCA 781  
QY 1809 TACAGTAT 1868  
DB 782 TACAGTAT 839  
QY 1869 AACTGTATAGAGACCAACAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1928  
DB 840 AACTGTATAGAGACCAACAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 899  
QY 1929 GAGAGAT 1988  
DB 900 GAGAGAT 959  
QY 1989 AGAATCATAGGCTTAT 2048  
DB 960 AGAATCATAGGCTTAT 1019  
QY 2049 CCTGTCTTGGAGAGATGATTTCTAGGTTCCATATAGGAAATTTCTTATTCAGAAA 2108  
DB 1020 CCTGTCTTGGAGAGATGATTTCTAGGTTCCATATAGGAAATTTCTTATTCAGAAA 1079  
QY 2109 GTCTGATAGGGCTTATAGCAAT 2168  
DB 1080 GTCTGATAGGGCTTATAGCAAT 1139  
QY 2169 GGCAGTGAAGTGAAGGACCAAGTATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2228  
DB 1140 GGCAGTGAAGTGAAGGACCAAGTATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1199  
QY 2229 ATTAATTTGGCTTTTGAATGATCTGATGCTGGAAGTGAAGGAAATCTTCAGAGCAT 2288  
DB 1200 ATTAATTTGGCTTTTGAATGATCTGATGCTGGAAGTGAAGGAAATCTTCAGAGCAT 1259  
QY 2289 GCTTATATTTGGGGCTTTGTCAGTATGGAACAGGACCTTGAACAGGAAAGCAATGTG 2348  
DB 1260 GCTTATATTTGGGGCTTTGTCAGTATGGAACAGGACCTTGAACAGGAAAGCAATGTG 1319  
QY 2349 ACTTAGCATGGGAATCAGGATTTTCTCTGAGGGCTATTTACAGGGTATATAGG 2408  
DB 1320 ACTTAGCATGGGAATCAGGATTTTCTCTGAGGGCTATTTACAGGGTATATAGG 1379  
QY 2409 TTTCAATCTTCAACAGGATATAGCAACAGGTTAACAGAAATCAATTCATTAATATCTA 2468  
DB 1380 TTTCAATCTTCAACAGGATATAGCAACAGGTTAACAGAAATCAATTCATTAATATCTA 1439  
QY 2469 AAACATGATCATATATATGATGATTTTCAATTTTCAATCTCAGGTTCCCTGA 2528  
DB 1440 AAACATGATCATATATATGATGATTTTCAATTTTCAATCTCAGGTTCCCTGA 1499  
QY 2529 TATGATTTCCATATACATGCTTCAATCCCTTTGTAATGATATCATATTTGGAAGG 2588  
DB 1500 TATGATTTCCATATACATGCTTCAATCCCTTTGTAATGATATCATATTTGGAAGG 1559  
QY 2589 CTATTTATATCTTATTTGCTGCTGACATGATGAGCCATGAGGCACTGTTATATATG 2648  
DB 1560 CTATTTATATCTTATTTGCTGCTGACATGATGAGCCATGAGGCACTGTTATATATG 1619  
QY 2649 AATGTCATCTCTGTTCAATCTTGAATGCTTTTGTCTATCTATGAAATCCCCAGAAAGT 2708  
DB 1620 AATGTCATCTCTGTTCAATCTTGAATGCTTTTGTCTATCTATGAAATCCCCAGAAAGT 1679  
QY 2709 GCTTAGAAT 2768  
DB 1680 GCTTAGAAT 1739  
QY 2769 TCTGTCTGGAACATATAGCCAGGCAATTTTCCAGCTTCTTTGAGTTGGGATATATATAA 2828  
DB 1740 TCTGTCTGGAACATATAGCCAGGCAATTTTCCAGCTTCTTTGAGTTGGGATATATATAA 1799

QY 2829 TTTCGGCCATTACTCTCCAAATGTAGTGAAGTGAATGCAATTTCTAATCTGCTCA 2888  
Db 1800 TTTCGGCCATTACTCTCCAAATGTAGTGAAGTGAATGCAATTTCTAATCTGCTCA 1859  
QY 2889 TAAACCTCCCATGTGAGGCTTCTCATGTGACATTAATGTGACCTGGGAAGCTATGT 2948  
Db 1860 TAAACCTCCCATGTGAGGCTTCTCATGTGACATTAATGTGACCTGGGAAGCTATGT 1919  
QY 2949 GTTACACAGAGTAATATCACAGGAAGCTGTGATTTCTGAAAAAATCTGACAGCCCAACC 3008  
Db 1920 GTTACACAGAGTAATATCACAGGAAGCTGTGATTTCTGAAAAAATCTGACAGCCCAACC 1979  
QY 3009 TCTGTCATTTGCACTCCCATCTTGTATTTGTAGACAGCACTGTGATTAAGTGAATTA 3068  
Db 1980 TCTGTCATTTGCACTCCCATCTTGTATTTGTAGACAGCACTGTGATTAAGTGAATTA 2039  
QY 3069 GTACTATTGTGTCAAG 3084  
Db 2040 GTACTATTGTGTCAAG 2055

RESULT 7  
AX267943 2061 bp DNA linear PAT 26-OCT-2001  
LOCUS Sequence 917 from Patent WO01/73032.  
DEFINITION AX267943  
ACCESSION AX267943.1 GI:16516537  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Kajos,M.D., Fanger,G.R., Reiter,M.W., Stolk,J.A., Day,C.H.,  
Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.  
and Henderson,R.A.  
Compositions and methods for the therapy and diagnosis of prostate  
cancer.  
Patent: WO 01/73032-A, 917 04-OCT-2001;  
JOURNAL CORIXA CORPORATION (US)  
TITLE Location/Qualifiers  
FEATURES  
source 1. 2061  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 612 a 383 c 399 g 667 t  
ORIGIN

Query Match 64.7%; Score 2028.4; DB 6; Length 2061;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1029 GATTGACAGAGGATCTTTCATTCATTCAGAGTCTGTGATTCAGATTTAATGTTAACT 1148  
Db 3 GATTGACAGAGGATCTTTCATTCATTCAGAGTCTGTGATTCAGATTTAATGTTAACT 122  
QY 1089 GTGAGTATCAAACTCTTTTCATTCAGAGTCTGTGATTCAGATTTAATGTTAACT 1148  
Db 63 GTGAGTATCAAACTCTTTTCATTCAGAGTCTGTGATTCAGATTTAATGTTAACT 122  
QY 1149 TTTGGAAGACAGATTCAGAAAAAATTTCTTAAATAAATACTGATGATCTTC 1208  
Db 123 TTTGGAAGACAGATTCAGAAAAAATTTCTTAAATAAATACTGATGATCTTC 181  
QY 1209 AAATATGAAGTGTGGGGAATCTCATTTTTCATATATTTTCTTCTTGTTCCT 1268  
Db 182 AAATATGAAGTGTGGGGAATCTCATTTTTCATATATTTTCTTCTTGTTCCT 241  
QY 1269 TGTATCATATATTTAATACCTTACAGTGTGTGTGAGGATTAATCTTTCAT 1328  
Db 242 TGTATCATATATTTAATACCTTACAGTGTGTGTGAGGATTAATCTTTCAT 301

QY 1329 TTTACATGACAGTCCAAATCTAAATCTGCTTCTACATGATGTTTACAGATTCAGATAA 1368  
Db 302 TTTACATGACAGTCCAAATCTAAATCTGCTTCTACATGATGTTTACAGATTCAGATAA 361  
QY 1389 GAATGTATCATCTAGAGAACTTTTGCCAAAGGCTTAAGCACAGGCAAGGAAATTAACAC 1448  
Db 362 GAATGTATCATCTAGAGAACTTTTGCCAAAGGCTTAAGCACAGGCAAGGAAATTAACAC 421  
QY 1449 AGAATATATTAATGAATATATCTAGCTTAAATCTAATCTTCTTCAAGACTCCCA 1508  
Db 422 AGAATATATTAATGAATATATCTAGCTTAAATCTAATCTTCTTCAAGACTCCCA 481  
QY 1509 ACCACATTGATCTTCAGAAAAATGCTGTCTTCAAAATGACTTTCACAGAGAAATTAAT 1568  
Db 482 ACCACATTGATCTTCAGAAAAATGCTGTCTTCAAAATGACTTTCACAGAGAAATTAAT 541  
QY 1569 TTTTCCTCTGAGACCTAGACCTTAAGGGGAAGATGGAAGTAAGCCTGAAAAGATAC 1628  
Db 542 TTTTCCTCTGAGACCTAGACCTTAAGGGGAAGATGGAAGTAAGCCTGAAAAGATAC 601  
QY 1629 ATTTACCTAGCTTAATGAATGACACACTGTCTGAGAGTTTTCACAGCATATGACC 1688  
Db 602 ATTTACCTAGCTTAATGAATGACACACTGTCTGAGAGTTTTCACAGCATATGACC 661  
QY 1689 CTGTTTTCCTATTTAATTTTCTTATCAACCTTTTAATTAGGCAAGATTAATTAATGAC 1748  
Db 662 CTGTTTTCCTATTTAATTTTCTTATCAACCTTTTAATTAGGCAAGATTAATTAATGAC 721  
QY 1749 CCTCATGCTAGCCATGAGGAAATGATGTTCAAGTGGGATCAAGTAATTAATGGGATCA 1808  
Db 722 CCTCATGCTAGCCATGAGGAAATGATGTTCAAGTGGGATCAAGTAATTAATGGGATCA 781  
QY 1809 TACAGATTAATAATTTAAAAAAGACTTATGCTCCCATCTCATATGATGTGAAG 1868  
Db 782 TACAGATTAATAATTTAAAAAAGACTTATGCTCCCATCTCATATGATGTGAAG 839  
QY 1869 AACTGTAGAGAGCCAAAGAGTGTGAGTGTAGAGATTTCCAGAGCTTACATTTCTA 1928  
Db 840 AACTGTAGAGAGCCAAAGAGTGTGAGTGTAGAGATTTCCAGAGCTTACATTTCTA 899  
QY 1929 GAGAGGATTTAATTTCTTCTCATCTATCATGATGTGTGAATTGGAATTTCTGGAAC 1988  
Db 900 GAGAGGATTTAATTTCTTCTCATCTATCATGATGTGTGAATTGGAATTTCTGGAAC 959  
QY 1989 AGAATCATGAGCTTAATCCACATGATTTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2048  
Db 960 AGAATCATGAGCTTAATCCACATGATTTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019  
QY 2049 CCTGTGTCTGGAAGAGTATTTCTAGCTTCCATTAATGGAAGATTTCTTATTCAGAAA 2108  
Db 1020 CCTGTGTCTGGAAGAGTATTTCTAGCTTCCATTAATGGAAGATTTCTTATTCAGAAA 1079  
QY 2109 GTCTGATAGGCTTAATGAGAGTATTTTAAATTTTAAATTTTCAATGATGATTTCTATA 2168  
Db 1080 GTCTGATAGGCTTAATGAGAGTATTTTAAATTTTAAATTTTCAATGATGATTTCTATA 1139  
QY 2169 GGCAGTATGTTAGGAGCAACAGTATATGGAAGTATGATGAGAGTCTTGAAG 2228  
Db 1140 GGCAGTATGTTAGGAGCAACAGTATATGGAAGTATGATGAGAGTCTTGAAG 1199  
QY 2229 ATTAATTTGCTTTTGAAGTGAATCTGTAGCTGGAAGTGAAGGAATCTTCAGAGCAT 2288  
Db 1200 ATTAATTTGCTTTTGAAGTGAATCTGTAGCTGGAAGTGAAGGAATCTTCAGAGCAT 1259  
QY 2289 GCTTATTTGGGCTTTGTCAGATGAGACAGGACCTTTCAGACAGGAAAGCAATCTG 2348  
Db 1260 GCTTATTTGGGCTTTGTCAGATGAGACAGGACCTTTCAGACAGGAAAGCAATCTG 1319  
QY 2349 ACTTAGCATGGAATCAGGCAATTTTGTCTTGAAGGCTATATTCACAGGTTAATAGG 2408  
Db 1320 ACTTAGCATGGAATCAGGCAATTTTGTCTTGAAGGCTATATTCACAGGTTAATAGG 1379

QY 2409 TTTCATCTTCAACAGATATGACAAAGGTTTACCAAGAAACCTCAATTAATAATA 2468  
 DB 1380 TTTCATCTTCAACAGATATGACAAAGGTTTACCAAGAAACCTCAATTAATAATA 1439  
 QY 2465 AAACATGTGATATATATGTGTAAAGTTTCATTTTCTTTTCAATCTCAGATCCCTGA 2528  
 DB 1440 AAACATGTGATATATATGTGTAAAGTTTCATTTTCTTTTCAATCTCAGATCCCTGA 1499  
 QY 2529 TATGATTCCTTAACATGATCTTCAATCCCTTTTGTATATGATATCAATTTTGAATATG 2588  
 DB 1500 TATGATTCCTTAACATGATCTTCAATCCCTTTTGTATATGATATCAATTTTGAATATG 1559  
 QY 2589 CTATTTAATATCTGTAATTTTGTCTGATCTGATTAAGCCATGAGGCACTGTTATATG 2648  
 DB 1560 CTATTTAATATCTGTAATTTTGTCTGATCTGATTAAGCCATGAGGCACTGTTATATG 1619  
 QY 2649 AATGTCATCTGTCATCATATGATGCTGCTTTTGTCTCATATGATATCCCAAGAAAGT 2708  
 DB 1620 AATGTCATCTGTCATCATATGATGCTGCTTTTGTCTCATATGATATCCCAAGAAAGT 1679  
 QY 2709 GGCCTGAACATATATAGTCTTATGCTTGACACCGGTTATTTTCAATCAACCTGATTCCT 2768  
 DB 1680 GGCCTGAACATATATAGTCTTATGCTTGACACCGGTTATTTTCAATCAACCTGATTCCT 1739  
 QY 2769 TCTGTCCTGAACATATAGTCTTATGCTTGACACCGGTTATTTTCAATCAACCTGATTCCT 2828  
 DB 1740 TCTGTCCTGAACATATAGTCTTATGCTTGACACCGGTTATTTTCAATCAACCTGATTCCT 1799  
 QY 2829 TCTGTCCTGAACATATAGTCTTATGCTTGACACCGGTTATTTTCAATCAACCTGATTCCT 2888  
 DB 1800 TCTGTCCTGAACATATAGTCTTATGCTTGACACCGGTTATTTTCAATCAACCTGATTCCT 1859  
 QY 2889 TAAACCCCTCCATGTGACGCTTTGATGATTAATATGATGATGATGATGATGATGATGAT 2948  
 DB 1860 TAAACCCCTCCATGTGACGCTTTGATGATTAATATGATGATGATGATGATGATGATGAT 1919  
 QY 2949 GTTACACAGATTAATACCAAGAGCCGATTTCTGAAAAAAGTGTGACAGGCAACACC 3008  
 DB 1920 GTTACACAGATTAATACCAAGAGCCGATTTCTGAAAAAAGTGTGACAGGCAACACC 1979  
 QY 3009 TCTGTCATTTGCAATCCCACTTGTATTTGTACAGGACAGTGTGATTAATTAATTAATTA 3068  
 DB 1980 TCTGTCATTTGCAATCCCACTTGTATTTGTACAGGACAGTGTGATTAATTAATTAATTA 2039  
 QY 3069 GTACATTTGTGCAAG 3084  
 DB 2040 GTACATTTGTGCAAG 2055  
 RESULT 8  
 BD160111  
 LOCUS 2061 bp DNA linear PAT 17-JAN-2003  
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION BD160111.1 GI:27865869  
 VERSION BD160111.1  
 KEYWORDS JP 2002191363-A/14954.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2061)  
 Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
 Primer for synthesizing full-length cDNA and use thereof  
 Patent: JP 2002191363-A, 14954 09-JUL-2002;  
 HELIX RESEARCH INSTITUTE  
 OS Homo sapiens (human)  
 PN JP 2002191363-A/14954  
 PD 09-JUL-2002  
 PF 28-JUL-2000 JP 2000280990  
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KIORU SAITO,  
 JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,

PI KEIICHI NAGAI, TETSUO OTSUKI  
 PC  
 C12N5/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N5/00, C12N5/00 CC  
 PC  
 Primer for synthesizing full-length cDNA and use thereof FH Key  
 FT source 1. .2061  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 612 a 383 c 399 g 667 t  
 ORIGIN  
 Query Match 64.7%; Score 2028.4; DB 6; Length 2061;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;  
 QY 1029 GATTGACAGCGCATCTTCGACTTTCCATGTCGACACAGCTTCGAGCCCTAGGT 1088  
 DB 3 GATTGACAGCGCATCTTCGACTTTCCATGTCGACACAGCTTCGAGCCCTAGGT 62  
 QY 1089 GTCAGATCAAACTTCTTCCATTCAGAGTCTCTGATTTGATTTTAATTTAATTAACAT 1148  
 DB 63 GTCAGATCAAACTTCTTCCATTCAGAGTCTCTGATTTGATTTTAATTTAATTTAATTAACAT 1122  
 QY 1149 TTTGAAAGCAGTATTCAGAAAAAATTTCTTAAATAAATAAATAAATAAATAAATAAATAA 1208  
 DB 123 TTTGAAAGCAGTATTCAGAAAAAATTTCTTAAATAAATAAATAAATAAATAAATAAATAA 181  
 QY 1209 AATAATGAACAGTATTCAGAAAAAATTTCTTAAATAAATAAATAAATAAATAAATAAATAA 1268  
 DB 182 AATAATGAACAGTATTCAGAAAAAATTTCTTAAATAAATAAATAAATAAATAAATAAATAA 241  
 QY 1269 TGTACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1328  
 DB 242 TGTACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 301  
 QY 1329 TTTACACAGCATCCCAATCTTAACTGTTCTGATGATGTTTACAGATTTCTGATTA 1388  
 DB 302 TTTACACAGCATCCCAATCTTAACTGTTCTGATGATGTTTACAGATTTCTGATTA 361  
 QY 1389 GATGATCATCTAGAGAACATTTGCCAAGGCTTACAGAGGCAAGGAAATTAACAC 1448  
 DB 362 GATGATCATCTAGAGAACATTTGCCAAGGCTTACAGAGGCAAGGAAATTAACAC 421  
 QY 1449 AGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1508  
 DB 422 AGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 481  
 QY 1509 ACCACATGATCTCAGAAAAATGCTGTCTCAAAATGACTTCAACAGAGAAATTAAT 1568  
 DB 482 ACCACATGATCTCAGAAAAATGCTGTCTCAAAATGACTTCAACAGAGAAATTAAT 541  
 QY 1569 TTTTCTCTGACACTAGCACTTAAGGGGAAGTTGGAAGTAAAGCTTGAAGAGATAC 1628  
 DB 542 TTTTCTCTGACACTAGCACTTAAGGGGAAGTTGGAAGTAAAGCTTGAAGAGATAC 601  
 QY 1629 ATTACATGATCTCAGAAAAATGCTGTCTCAAAATGACTTCAACAGAGAAATTAAT 1688  
 DB 602 ATTACATGATCTCAGAAAAATGCTGTCTCAAAATGACTTCAACAGAGAAATTAAT 661  
 QY 1689 CTGTTTCTCTATTAATTTTCTATCAACCTTTAATTAAGCAAGATTAATTAATTAATTA 1748  
 DB 662 CTGTTTCTCTATTAATTTTCTATCAACCTTTAATTAAGCAAGATTAATTAATTAATTA 721  
 QY 1749 CCGATGTCAGCACTGGAAGAAATGATGTTGAGGGGATTCAGAGATTAATTAATTAATTAAT 1808  
 DB 722 CCGATGTCAGCACTGGAAGAAATGATGTTGAGGGGATTCAGAGATTAATTAATTAATTAATTA 781





QY 1149 TTGGAAGACAGATATCAGAAAAAATTTCTTATATAAAAAATCAATCAGATCTTC 1208  
 Db 123 TTGGAAAGACAGATATCAGAAAAAATTTCTTAT -AAAAATCAATCAGATCTTC 181  
 QY 1209 AATATGAAGCTGGTGGGAATCTCCATTTTTCATATATTTCTTTCTTTCTTCT 1268  
 Db 182 AATATGAAGCTGGTGGGAATCTCCATTTTTCATATATTTCTTTCTTTCTTCT 241  
 QY 1269 TGTACATATATATATATATACCTGACTAGTGGTGGTGGAGGATTAATCTTTGAT 1328  
 Db 242 TGTACATATATATATATATATACCTGACTAGTGGTGGTGGAGGATTAATCTTTGAT 301  
 QY 1329 TTTACATGAGCTCCAAATCTAACTGCTTCTACTGATGCTTTACAGCATTTGAGATTA 1388  
 Db 302 TTTACATGAGCTCCAAATCTAACTGCTTCTACTGATGCTTTACAGCATTTGAGATTA 361  
 QY 1389 GAATGATCATCTAGAGAACATTTGGCAAGGCTTAGAGCGGCAAGGAAATTAACAC 1448  
 Db 362 GAATGATCATCTAGAGAACATTTGGCAAGGCTTAGAGCGGCAAGGAAATTAACAC 421  
 QY 1449 AGAATTAATTAATATAGATTAATCTAGCTTAAACTTAATCTTCTTCTAGAACTCCA 1508  
 Db 422 AGAATTAATTAATATAGATTAATCTAGCTTAAACTTAATCTTCTTCTAGAACTCCA 481  
 QY 1509 ACCAGATGAGATCTGAGAAAAATGCTGCTCAAAATGACTTCAAGAGAAATTAAT 1568  
 Db 482 ACCAGATGAGATCTGAGAAAAATGCTGCTCAAAATGACTTCAAGAGAAATTAAT 541  
 QY 1569 TTTTCTCTGACACTAGCACTTAAAGGGAGAGATTGGAATTAAGCTTGAAGAGATAC 1628  
 Db 542 TTTTCTCTGACACTAGCACTTAAAGGGAGAGATTGGAATTAAGCTTGAAGAGATAC 601  
 QY 1629 ATTATCTAGTAAATGAAGTGAACAAGTGTCTGAGAGATTTTCAAGATATAGACC 1688  
 Db 602 ATTATCTAGTAAATGAAGTGAACAAGTGTCTGAGAGATTTTCAAGATATAGACC 661  
 QY 1689 CTGTTTCTCTATTTAATTTCTTATCAACCTTTAATTTAGGCAAGATTAATATAGAC 1748  
 Db 662 CTGTTTCTCTATTTAATTTCTTATCAACCTTTAATTTAGGCAAGATTAATATAGAC 721  
 QY 1749 CCMCATTTGAGCAATGGGAAATTTGATGTTCACTGGGATCAGTAAATTAATGGGGTCA 1808  
 Db 722 CCMCATTTGAGCAATGGGAAATTTGATGTTCACTGGGATCAGTAAATTAATGGGGTCA 781  
 QY 1809 TACAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1868  
 Db 782 TACAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 839  
 QY 1865 AACTGTTAGAGAACCAACAGGGTATGAGGTTAGAGATTTCCAGATCTTACATTTCTA 1928  
 Db 840 AACTGTTAGAGAACCAACAGGGTATGAGGTTAGAGATTTCCAGATCTTACATTTCTA 899  
 QY 1929 GAGGAGATTAATTTCTTCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1988  
 Db 900 GAGGAGATTAATTTCTTCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 959  
 QY 1989 ABAATCTAGCTTTATATCCACTAGCTATGCTTATGCTGCTGCTCAATTTCAATTA 2048  
 Db 960 ABAATCTAGCTTTATATCCACTAGCTATGCTTATGCTGCTGCTCAATTTCAATTA 2019  
 QY 2049 CCGTGTCTTGGAAAGAGATTTCTAGGTTCAACATTAAGAAATCTTATTCAGAAA 2108  
 Db 1020 CCGTGTCTTGGAAAGAGATTTCTAGGTTCAACATTAAGAAATCTTATTCAGAAA 1079  
 QY 2109 GTCTGATAGGGCTTATAGCAAGTATTAATTTTAAAGTTCATAGGATCTGATA 2168  
 Db 1080 GTCTGATAGGGCTTATAGCAAGTATTAATTTTAAAGTTCATAGGATCTGATA 1139  
 QY 2169 GGCAGTAGCTTATAGGAGCCACCACTTATGATGGAAGTATGAGATGCGAGCTTGAAG 2228  
 Db 1140 GGCAGTAGCTTATAGGAGCCACCACTTATGATGGAAGTATGAGATGCGAGCTTGAAG 1199

QY 2229 ATACATTTGGCTTTTGAAGTGAATCTGATGTAAGGAAAGTGAAGGATCTTCAGACCAT 2288  
 Db 1200 ATACATTTGGCTTTTGAAGTGAATCTGATGTAAGGAAAGTGAAGGATCTTCAGACCAT 1259  
 QY 2289 GCTTTATTTGGGCTTTTGAAGTGAATCTGATGTAAGGAAAGTGAAGGATCTTCAGACCAT 2348  
 Db 1260 GCTTTATTTGGGCTTTTGAAGTGAATCTGATGTAAGGAAAGTGAAGGATCTTCAGACCAT 1319  
 QY 2349 ACTTAGGATGGAATCAGGCAATTTTGTCTGAGAGGCTATTAACAGGCTTAATAGG 2408  
 Db 1320 ACTTAGGATGGAATCAGGCAATTTTGTCTGAGAGGCTATTAACAGGCTTAATAGG 1379  
 QY 2409 TTTTATCTTCAAGAGATATGACACAGTGTAAACCAAGAACTCAAAATTAATTAATTA 2468  
 Db 1380 TTTTATCTTCAAGAGATATGACACAGTGTAAACCAAGAACTCAAAATTAATTAATTA 1439  
 QY 2469 AAACATGATATCATATATGATGTAAGTGAATCTTTTCTTCAATCTCAGGTTCCCTGA 2528  
 Db 1440 AAACATGATATCATATATGATGTAAGTGAATCTTTTCTTCAATCTCAGGTTCCCTGA 1499  
 QY 2529 TATGATTTCTATTAACATGCTTTCATCCCTTTTGTATGATATCATATTTGAAATGC 2588  
 Db 1500 TATGATTTCTATTAACATGCTTTCATCCCTTTTGTATGATATCATATTTGAAATGC 1559  
 QY 2589 CTATTTAATCTTGTATTTTGTCTGAGTGAATGAGGCTTGAAGGCTTGAATTTATATG 2648  
 Db 1560 CTATTTAATCTTGTATTTTGTCTGAGTGAATGAGGCTTGAAGGCTTGAATTTATATG 1619  
 QY 2649 AATGATCTCTGATCATATGATGCTCTTGTGATCATATTAATCCCGACGAAAGT 2708  
 Db 1620 AATGATCTCTGATCATATGATGCTCTTGTGATCATATTAATCCCGACGAAAGT 1679  
 QY 2709 GCTTAAAGCAATATAGTGTATGCTTGAACCGGATATTTTTCATCAACCGATTCCT 2768  
 Db 1680 GCTTAAAGCAATATAGTGTATGCTTGAACCGGATATTTTTCATCAACCGATTCCT 1739  
 QY 2769 TCTGCTGTAACACATAGGCAAGCAATTTTCCAGCTTCTTGAATGAGGATTAATTA 2828  
 Db 1740 TCTGCTGTAACACATAGGCAAGCAATTTTCCAGCTTCTTGAATGAGGATTAATTA 1799  
 QY 2829 TTTGAGCATTAATTTTCAATGATGAGGAAAGTGAATTTCTATTAATCTGCTCA 2888  
 Db 1800 TTTGAGCATTAATTTTCAATGATGAGGAAAGTGAATTTCTATTAATCTGCTCA 1859  
 QY 2889 TAAACCTTCCCATGAGCAAGCTTTCATGTTGATTAATGATGTTGAGGAAAGTGAAT 2948  
 Db 1860 TAAACCTTCCCATGAGCAAGCTTTCATGTTGATTAATGATGTTGAGGAAAGTGAAT 1919  
 QY 2949 GTTACACAGATTAATCACCAGAGGCTTGAATTTTGAATAAATCTGCAAGCCAAAC 3008  
 Db 1920 GTTACACAGATTAATCACCAGAGGCTTGAATTTTGAATAAATCTGCAAGCCAAAC 1979  
 QY 3009 TCTGATTTGCAATCCCACTGCTTATTTGTAAGGAGGCTTGAATTAATTAATTA 3068  
 Db 1980 TCTGATTTGCAATCCCACTGCTTATTTGTAAGGAGGCTTGAATTAATTAATTA 2039  
 QY 3069 GATATATGTTCAAG 3084  
 Db 2040 GATATATGTTCAAG 2055

RESULT 10  
 AC025249/c 179664 bp DNA linear HTG 01-SEP-2000  
 LOCUS Homo sapiens chromosome 11 clone RP11-648B16, WORKING DRAFT  
 DEFINITION  
 ACCESSION AC025249  
 VERSION AC025249.5 GI:9958307  
 KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.



REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1	(bases 1 to 175664)	Waterston, R.H.	The sequence of Homo sapiens clone	2	(bases 1 to 175664)	Waterston, R.H.	Direct Substans
			Unpublished				Submitted (07-MAR-2000) Genome Seq

COMMENT On Sep 1, 2000 this sequence version replaced gi:7631097.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 conigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the conigs are represented as  
\* runs of 'N', but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

*	1	25906:	contig of 25906 bp in length
*	25907	26006:	gap of unknown length
*	26007	63431:	contig of 37425 bp in length
*	63432	63531:	gap of unknown length
*	63532	126351:	contig of 7828 bp in length
*	136360	136459:	gap of unknown length
*	136460	141116:	contig of 4657 bp in length
*	141117	141216:	gap of unknown length
*	141217	145465:	contig of 4253 bp in length
*	145470	145569:	gap of unknown length
*	145570	154744:	contig of 9174 bp in length
*	154744	154843:	gap of unknown length
*	154844	155972:	contig of 1130 bp in length
*	155972	166073:	gap of unknown length
*	166074	179666:	contig of 13591 bp in length

FEATURES	Location/Qualifiers
source	1. .179664

```
misc_feature      1. .25506  
/note="assembly_name:Contig10  
26007. .63431  
/note="assembly_name:Contig11"  
6352. .16359  
/note="assembly_name:Contig12  
136660. .14116  
/note="assembly_name:Contig15"  
14117. .145469  
/note="assembly_name:Contig16"  
145570. .154743
```

```

misc_feature      /note="assembly_name:Contig7
                   clone_end:T7
                   vector_side:right"
                   154844 .165973
/note="assembly_name:Contig98"
misc_feature      166074 .179664
                   /note="assembly_name:Contig99
                   clone_end:SP6
                   vector_side:right"
BASE COUNT      51388 a 36250 g 53893 t 700 others
ORIGIN

```

OY	1117	GAGCTCCGATTCAGATTATTAATGTAAACATTTGGAAAGCAGTATTCAGAAAAAAT	1176
Db	165973	GAGCTCCGATTCAGATTATTAATGTAAACATTTGGAAAGCAGTATTCAGAAAAAAT	16591
OY	1177	TTCCTTAATAAAAATCAACTCAGATCCTTCAAATGAACCTGGTGGGAAATCTTCCA	1236
Db	165913	TTCCTTAAT-AAAAATACAACAGATCCTTCAAATGAACCTGGTGGGAAATCTTCCA	165855
OY	1237	TTTTTTCATATTAATTTCTTGTTGTTCTGTGCATCATTAATTAATTAATACCCTGAC	1296
Db	165854	TTTTTTCATATTAATTTCTTGTTGTTCTGTGCATCATTAATTAATTAATACCCTGAC	16579
OY	1297	TAGGTTGTGGTGGAGGTTATTACTTTCAATTTTACATGACGTCCAATCTAACTGC	1356
Db	165794	TAGGTTGTGGTGGAGGTTATTACTTTCAATTTTACATGACGTCCAATCTAACTGC	165735
OY	1357	TTCTACTGATGGTTTACAGCAATTCGAGATAGAATGGTACATCTAGAGAACATTTGCCA	1416
Db	165734	TTCTACTGATGGTTTACAGCAATTCGAGATAGAATGGTACATCTAGAGAACATTTGCCA	165675
OY	1417	AAGCCTTAGCACCGGCAAAAGAAAATTAACAACAGAAATTAATTAATGAATTAATCTAGC	1476
Db	165674	AAGCCTTAGCACCGGCAAAAGAAAATTAACAACAGAAATTAATTAATGAATTAATCTAGC	165615
OY	1477	TTTAAACTTAATCTCTCTTCAGAACTCCCAACCAATGGATCTCAGAAAAATGCTGT	1536
Db	165614	TTTAAACTTAATCTCTCTTCAGAACTCCCAACCAATGGATCTCAGAAAAATGCTGT	165555
OY	1537	CTTCAAAATGACTTCTACAGAGAGAAATTAATTTTCTCTGCACATGACATTTAAGGG	1596
Db	165554	CTTCAAAATGACTTCTACAGAGAGAAATTAATTTTCTCTGCACATGACATTTAAGGG	165495
OY	1597	GAGATTTGGAATGAAGCCTTGAAGAAAGTCACTTTCACCTACGTATGAAGAGTTGCAC	1656
Db	165494	GAGATTTGGAATGAAGCCTTGAAGAAAGTCACTTTCACCTACGTATGAAGAGTTGCAC	165433
OY	1657	ACTGTTCTGAGAGTTTTCACAGCATATGAGACCCGTGTTTTCATATTAATTTCTTATCA	1716
Db	165434	ACTGTTCTGAGAGTTTTCACAGCATATGAGACCCGTGTTTTCATATTAATTTCTTATCA	165375
OY	1717	ACCCTTTAATTAGGCMAAGATATTATTAGTACCCCTCATTTGAGCATGGGAAATTGATG	1776
Db	165374	ACCCTTTAATTAGGCMAAGATATTATTAGTACCCCTCATTTGAGCATGGGAAATTGATG	165315
OY	1777	TTTCACTGGGGAATCACTGAATTAATGGGGTATTCACAGTATTAATAAAAAAAAAA	1836
Db	165314	TTTCACTGGGGAATCACTGAATTAATGGGGTATTCACAGTATTAATAAAAAAAAAA	165257
OY	1837	GACTTCATGCCCMAATCTCATATGATGTGGAAAGAACTGTTAGAGAGCAACAAGGTATG	1896
Db	165256	GACTTCATGCCCMAATCTCATATGATGTGGAAAGAACTGTTAGAGAGCAACAAGGTATG	165197
OY	1897	GGTTGAGATTTCCAGAGCTTCAATTTCTAGAGAGAGATTAATTTCTTCTCACTCA	1956
Db	165196	GGTTGAGATTTCCAGAGCTTCAATTTCTAGAGAGAGATTAATTTCTTCTCACTCA	165137
OY	1957	TCCAAGTGTGATTTAGAAATTTCTGGCAACAAGAACTCAATGSCCTTTAATCCCACTAGCT	2016

Db	165136	TTCCAGGCTGTATTTAGGATTTCTGGCAACAGAACTAGGCTTTATCCACAGT	165077
Qy	2017	ATTGCTATATGCTGGTCCAAATTCGCAATTAACCTGCTGTTGGAAAGTATTTCTAG	2076
Db	165076	ATTGCTATATGCTGGTCCAAATTCGCAATTAACCTGCTGTTGGAAAGTATTTCTAG	165017
Qy	2077	GTTTACCATTTAGGAAGATTTCTATTCAGAAAGTGCATAGGCTTTATGCAAGTATT	2136
Db	165016	GTTTACCATTTAGGAAGATTTCTATTCAGAAAGTGCATAGGCTTTATGCAAGTATT	164957
Qy	2137	TATTTTAAAGTTCATAGGATTTCTATTCAGAAAGTGCATAGGCTTTATGCAAGTATT	2196
Db	164956	TATTTTAAAGTTCATAGGATTTCTATTCAGAAAGTGCATAGGCTTTATGCAAGTATT	164897
Qy	2197	TGATGGAGATAGATAGGAGCTCTGTAAGATAATTCATAGGCTTTATGCAAGTATT	2256
Db	164896	TGATGGAGATAGATAGGAGCTCTGTAAGATAATTCATAGGCTTTATGCAAGTATT	164837
Qy	2257	TAGCTGGAAGTAGGGAATCTTCAGAGCAATGCTTTATTTGAGGCTTTGTCAGATGG	2316
Db	164836	TAGCTGGAAGTAGGGAATCTTCAGAGCAATGCTTTATTTGAGGCTTTGTCAGATGG	164777
Qy	2317	AACAGGAACTTTGAGACCAAGAAAGCAATCTGCTAGGCAATGGAATCAGGATTTTG	2376
Db	164776	AACAGGAACTTTGAGACCAAGAAAGCAATCTGCTAGGCAATGGAATCAGGATTTTG	164717
Qy	2377	CTTCTGAGGGGCTATTACCAAGGGTTAATAGGTTTCATCTTCAACAGATATGACACAG	2436
Db	164716	CTTCTGAGGGGCTATTACCAAGGGTTAATAGGTTTCATCTTCAACAGATATGACACAG	164657
Qy	2437	TGTTAACCAAGAACTCAATTTACAAATCTAAACATGTAATGCTATATGCTGTAATT	2496
Db	164656	TGTTAACCAAGAACTCAATTTACAAATCTAAACATGTAATGCTGTAATT	164597
Qy	2497	TCAATTTCTTTTCAATCTCAGGTTCCCTGATATGATTTCTATATGATGCTTTATCC	2556
Db	164596	TCAATTTCTTTTCAATCTCAGGTTCCCTGATATGATTTCTATATGATGCTTTATCC	164537
Qy	2557	CCTTTGTATGATATCAATTTTGGAAATGCTTTTAAATCTGTAATTTGCTGCTGGA	2616
Db	164536	CCTTTGTATGATATCAATTTTGGAAATGCTTTTAAATCTGTAATTTGCTGCTGGA	164477
Qy	2617	CTGTAAAGCCATAGGAGCACTGTTATTAATGATGATCATCTGCTTATGATGACTGC	2676
Db	164476	CTGTAAAGCCATAGGAGCACTGTTATTAATGATGATCATCTGCTTATGATGACTGC	164417
Qy	2677	TCTTGTCTATCATTTGAATCCCCAGCAAGTGCCTAGAACATAATAGTCTTATGCTTG	2736
Db	164416	TCTTGTCTATCATTTGAATCCCCAGCAAGTGCCTAGAACATAATAGTCTTATGCTTG	164357
Qy	2737	ACAACGGTATTTTTCATCAAACTGATTCCTTGTCTGTAACATAGGCAAGCAAT	2796
Db	164356	ACAACGGTATTTTTCATCAAACTGATTCCTTGTCTGTAACATAGGCAAGCAAT	164297
Qy	2797	TTCCAGGCTCTTTAGTGGTATTTAAATTTCTGGCCATTTACTTCAATGATGAGTG	2856
Db	164296	TTCCAGGCTCTTTAGTGGTATTTAAATTTCTGGCCATTTACTTCAATGATGAGTG	164237
Qy	2857	AAGTACATGTGCAATTTCTATACCTGGCTATAAACTCCCTGCTGAGGCTTTTCAAT	2916
Db	164236	AAGTACATGTGCAATTTCTATACCTGGCTATAAACTCCCTGCTGAGGCTTTTCAAT	164177
Qy	2917	GTTGACATTAATGTGACTTGGGAAGCTATGTGTTACACAGATTAATCACCAAGGCT	2976
Db	164176	GTTGACATTAATGTGACTTGGGAAGCTATGTGTTACACAGATTAATCACCAAGGCT	164117
Qy	2977	GGATTTCTGAAAAAATCTGTGAGAGCAAACTCTGTCAATTTGCAACTCCCACTTGATT	3036
Db	164116	GGATTTCTGAAAAAATCTGTGAGAGCAAACTCTGTCAATTTGCAACTCCCACTTGATT	164057
Qy	3037	TGTACAGAGCACTTGATTAAGTAAAAATTAAGTATGTTGTGCAAGAAAAAAA	3094
Db	164056	TGTACAGAGCACTTGATTAAGTAAAAATTAAGTATGTTGTGCAAGTCTTGAAAA	163995

Db	164056	TGTACAGAGCACTTGATTAAGTAAAAATTAAGTATGTTGTGCAAGTCTTGAAAA	163995
RESULT 11			
AX201143			
LOCUS	AX201143	1302 bp	DNA
DEFINITION	Sequence 773 from Patent WO0151633.		
ACCESSION	AX201143		
VERSION	AX201143.1	GI:15390907	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Xu, J., Dillon, P. C., Mitcham, J. L., Hatlock, S. L., Jiang, Y., Reed, S. G., Kalos, M. D., Fanger, G. R., Day, C. H., Reiter, M. W., Stolk, J. A., Skelky, Y. A., Wang, A. and Meagher, M. J.		
TITLE	Compositions and methods for the therapy and diagnosis of prostate cancer.		
JOURNAL	Patent: WO 0151633-A 773 19-JUL-2001; CORIAX CORPORATION (US)		
FEATURES			
SOURCE	Location/Qualifiers		
	1..1302		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	277 a 346 c 259 g 420 t		
ORIGIN			
Query Match	41.1%; Score 1287.8; DB 6; Length 1302;		
Best Local Similarity	99.8%; Pred. No. 2,1e-274;		
Matches 1300; Conservative	0; Mismatches 2; Indels 1; Gaps 1;		
Qy	52	TGACAAAGGGGCTACACATTCCTCCATACAGGTGAGGCTCTACCTGCTGGCTGG	111
Db	1	TGGCAAGGGGCTACACATTCCTCCATACAGGTGAGGCTCTACCTGCTGGCTGG	60
Qy	112	TCAAGTTCAAGCTTTCTTATGATGTGATGCCAATGSCAATGATCCAGTGTACATAC	171
Db	61	TCAAGTTCAAGCTTTCTTATGATGTGATGCCAATGSCAATGATCCAGTGTACATAC	120
Qy	172	TTATCCTTAATAGGCTCCCTGTTAGAAAGAGCTCAGTTCGCTGGCTTCCCATG	231
Db	121	TTATCCTTAATAGGCTCCCTGTTAGAAAGAGCTCAGTTCGCTGGCTTCCCATG	180
Qy	232	TGCTCCCTTACTTATTTGCTGTGTAAGTAACTTGACATCATCTAATTTGCGAGCT	291
Db	181	TGCTCCCTTACTTATTTGCTGTGTAAGTAACTTGACATCATCTAATTTGCGAGCT	240
Qy	292	GAGCAGAGCTGATAGAGCCATGATATATTTCTTTCATGCTTTCAGGCAATGACATC	351
Db	241	GAGCAGAGCTGATAGAGCCATGATATATTTCTTTCATGCTTTCAGGCAATGACATC	300
Qy	352	CTCATCTTCACTCATCATGTCACAAATGCTGGCCATCTTGTGTTCAATTCATACC	411
Db	301	CTCATCTTCACTCATCATGTCACAAATGCTGGCCATCTTGTGTTCAATTCATACC	360
Qy	412	ATCCAGTTGATGCTGTGCTGCTACAGATTTTTCATGCAATCCTTATCTAGGCAATG	471
Db	361	ATCCAGTTGATGCTGTGCTGCTACAGATTTTTCATGCAATCCTTATCTAGGCAATG	420
Qy	472	TCCACAGTCTGTGCGCATGAGCTTTTGAACCGCTATGTGAGCATGTGACCACTGCG	531
Db	421	TCCACAGTCTGTGCGCATGAGCTTTTGAACCGCTATGTGAGCATGTGACCACTGCG	480
Qy	532	CATGCAAGTACTTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	591
Db	481	CATGCAAGTACTTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540
Qy	592	GGGCTGCACTGATGCAACCCCTTCTGTCTTATCAAGAGCTGCTGCTGCTGCTGCTG	651
Db	541	GGGCTGCACTGATGCAACCCCTTCTGTCTTATCAAGAGCTGCTGCTGCTGCTGCTG	600

QY 652 AATATCCCTTCCCATCTCTACTGCTACACCAAGATGTCATGAGCTGCTGATGAT 711  
 DB 601 AATATCCCTTCCCATCTCTACTGCTACACCAAGATGTCATGAGCTGCTGATGAT 660  
 QY 712 ATCCGGGTCAATGTCCTTATGAGCTTATGCTATATCTCCGCCATTGGCTGACTCA 771  
 DB 661 ATCCGGGTCAATGTCCTTATGAGCTTATGCTATATCTCCGCCATTGGCTGACTCA 720  
 QY 772 CTTCATCTCTCTCATATCTGCTTATCTTAAAGCTGTCGAGCTTGAAGTGA 831  
 DB 721 CTTCATCTCTCTCATATCTGCTTATCTTAAAGCTGTCGAGCTTGAAGTGA 780  
 QY 832 GCCCAGGCCAAGGCATTTGGCACTTGGCTCTCATGTCGTCGTCGTCATTTCTAT 891  
 DB 781 GCCCAGGCCAAGGCATTTGGCACTTGGCTCTCATGTCGTCGTCGTCATTTCTAT 840  
 QY 892 GTACCTTTCAATGATGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 951  
 DB 841 GTACCTTTCAATGATGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 900  
 QY 952 CCGCTCATCTTGGCCATATCTATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1011  
 DB 901 CCGCTCATCTTGGCCATATCTATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 960  
 QY 1012 GAGATGAAGCAAGAGATTCAGACGCGCATCTTCGACTTTTCATGTCGTCGTCGTCG 1071  
 DB 961 GAGATGAAGCAAGAGATTCAGACGCGCATCTTCGACTTTTCATGTCGTCGTCGTCG 1020  
 QY 1072 GATTCAGAGCCCTAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1131  
 DB 1021 GATTCAGAGCCCTAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1080  
 QY 1132 GATTTTATGTTAACTTTTGAAGACAGTATTCAGAAAAAAATTTCTTATTAATAA 1191  
 DB 1081 GATTTTATGTTAACTTTTGAAGACAGTATTCAGAAAAAAATTTCTTATTAATAA 1139  
 QY 1192 TACACCTCAGATCTCTCAATATGAAGTGGTGGGAAATCTCCATTTTTCATATATAT 1251  
 DB 1140 TACACCTCAGATCTCTCAATATGAAGTGGTGGGAAATCTCCATTTTTCATATATAT 1199  
 QY 1252 TTTCTTCTTGTGTTCTTGTCTACATATATATTAATTAATTAATTAATTAATTAAT 1311  
 DB 1200 TTTCTTCTTGTGTTCTTGTCTACATATATTAATTAATTAATTAATTAATTAAT 1259  
 QY 1312 GGGTTATTAATTTTCAATTTTACATGAGTCCAAATCTAAACT 1354  
 DB 1260 GGGTTATTAATTTTCAATTTTACATGAGTCCAAATCTAAACT 1302

RESULT 12  
 AX267942 1302 bp DNA linear PAT 26-OCT-2001  
 LOCUS AX267942  
 DEFINITION Sequence 916 from Patent WO0173032.  
 ACCESSION AX267942  
 VERSION AX267942.1 GI:16516536  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Xu, J., Dillon, D.C., Mitcham, J.T., Harlocker, S.T., Jiang, Y.,  
 Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H.,  
 Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skelky, Y.A., Hepler, W.T.  
 and Henderson, R.A.  
 TITLE Compositions and methods for the therapy and diagnosis of prostate  
 cancer. WO 0173032-A 916 04-OCT-2001;  
 JOURNAL CORIXA CORPORATION (US)  
 FEATURES  
 1. location/Qualifiers  
 1.1302  
 /organism="Homo sapiens"

BASE COUNT 277 a 346 c 259 g 420 t  
 ORIGIN  
 Query Match 41.1%; Score 1287.8; DB 6; Length 1302;  
 Best Local Similarity 99.8%; Pred. No. 2.1e-274;  
 Matches 1300; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 52 TGAACAAAGGGGGGACACAAATTCCTTCATACGGTTGAGCTGCTACCTGCTGGTGG 111  
 DB 1 TGAACAAAGGGGGGACACAAATTCCTTCATACGGTTGAGCTGCTACCTGCTGGTGG 60  
 QY 112 TCACAGTTACGCTTTCATGATGTCGATGTCATGTCATGTCATGTCATGTCATGTCAT 171  
 DB 61 TCACAGTTACGCTTTCATGATGTCGATGTCATGTCATGTCATGTCATGTCATGTCAT 120  
 QY 172 TTTCACTCTAATAGGCTCCCTGCTGTTAGAGAGCTCAATTTGCTGCTGCTCCCATGG 231  
 DB 121 TTTCACTCTAATAGGCTCCCTGCTGTTAGAGAGCTCAATTTGCTGCTGCTCCCATGG 180  
 QY 232 TGCTCCCTTACCTTATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 291  
 DB 181 TGCTCCCTTACCTTATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 240  
 QY 292 GAGCAGAGCTGTCATGAGCCCATGTCATATATTTCTTGCATGCTTTAGGTCATGTCAT 351  
 DB 241 GAGCAGAGCTGTCATGAGCCCATGTCATATATTTCTTGCATGCTTTAGGTCATGTCAT 300  
 QY 352 CTCATCTCCACCTCATCCATGAGCCCAAAATCTGGCCATCTTCTGTTCAATTCACATACC 411  
 DB 301 CTCATCTCCACCTCATCCATGAGCCCAAAATCTGGCCATCTTCTGTTCAATTCACATACC 360  
 QY 412 ATCCAGTTGATGCTTGTCTGTCACAGATTTTTCATCCATCTCTTATCTGGCATGGA 471  
 DB 361 ATCCAGTTGATGCTTGTCTGTCACAGATTTTTCATCCATCTCTTATCTGGCATGGA 420  
 QY 472 TCCACAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 531  
 DB 421 TCCACAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 480  
 QY 532 CATGCCACAGTACTTACGTTGCTCTGTCACCAAAATGTCGTCGTCGTCGTCGTCGTCG 591  
 DB 481 CATGCCACAGTACTTACGTTGCTCTGTCACCAAAATGTCGTCGTCGTCGTCGTCGTCGTCG 540  
 QY 592 GGGGTGCACTGATGAGCAACCCCTTCTGCTTATCATGAGAGCTGCCCTTCGCGGCTCC 651  
 DB 541 GGGGTGCACTGATGAGCAACCCCTTCTGCTTATCATGAGAGCTGCCCTTCGCGGCTCC 600  
 QY 652 AATATCCCTTCCCATCTCTACTGCTACACCAAGATGTCATGAGCTGAGCTGATGAT 711  
 DB 601 AATATCCCTTCCCATCTCTACTGCTACACCAAGATGTCATGAGCTGAGCTGATGAT 660  
 QY 712 ATCCGGGTCAATGTCCTTATGAGCTTATGCTATATCTCCGCCATTGGCTGACTCA 771  
 DB 661 ATCCGGGTCAATGTCCTTATGAGCTTATGCTATATCTCCGCCATTGGCTGACTCA 720  
 QY 772 CTTCATCTCTCTCATATCTGCTTATCTTAAAGCTGTCGAGCTTGAAGTGA 831  
 DB 721 CTTCATCTCTCTCATATCTGCTTATCTTAAAGCTGTCGAGCTTGAAGTGA 780  
 QY 832 GCCCAGGCCAAGGCATTTGGCACTTGGCTCTCATGTCGTCGTCGTCGTCATTTCTAT 891  
 DB 781 GCCCAGGCCAAGGCATTTGGCACTTGGCTCTCATGTCGTCGTCGTCGTCATTTCTAT 840  
 QY 892 GTACCTTTCAATGATGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 951  
 DB 841 GTACCTTTCAATGATGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 900  
 QY 952 CCGCTCATCTTGGCCATATCTATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1011  
 DB 901 CCGCTCATCTTGGCCATATCTATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 960

QY 1012 GGAGTGAAGACAAAGAGATGTGACAGCCATCTTCACTTTTCATGTGGCCACAC 1071  
 DB 961 GGAGTGAAGACAAAGAGATGTGACAGCCATCTTCACTTTTCATGTGGCCACAC 1020  
 QY 1072 GCTTCAGAGCCCTAGGTGTCAAGATCAAACTCTTTTCATGTAGAGCCCTGATTC 1131  
 DB 1021 GCTTCAGAGCCCTAGGTGTCAAGATCAAACTCTTTTCATGTAGAGCCCTGATTC 1080  
 QY 1132 GATTTTAAGTTAACTTTTGAAGACAGATTCAGAAAAAATTTCTTAATAAAA 1191  
 DB 1081 GATTTTAAGTTAACTTTTGAAGACAGATTCAGAAAAAATTTCTTAATAAAA 1139  
 QY 1192 TCAACTAGATCTTCAATATGAAGTGTGGAGATCCCTTTTCATATATAT 1251  
 DB 1140 TCAACTAGATCTTCAATATGAAGTGTGGAGATCCCTTTTCATATATAT 1199  
 QY 1252 TTTCTTTCTTTCTTTCTTCAATATATATATATATATATATATATATATAT 1311  
 DB 1200 TTTCTTTCTTTCTTTCTTCAATATATATATATATATATATATATATATAT 1259  
 QY 1312 GGGTTATCTTTTCAATTTTACATGAGTCCCAATCTTAAT 1354  
 DB 1260 GGGTTATCTTTTCAATTTTACATGAGTCCCAATCTTAAT 1302

RESULT 13  
 BC022401 1225 bp mRNA linear PRI 04-FEB-2002  
 LOCUS Homo sapiens, clone MGC:24137 IMAGE:4694856, mRNA, complete cds.  
 DEFINITION BC022401.1 GI:18490239  
 ACCESSION MGC.  
 VERSION BC022401.1 GI:18490239  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1225)  
 Strausberg, R.  
 Direct Submission  
 Submitted (01-FEB-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: CLONTECH  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: http://www.shgc.stanford.edu  
 Contact: (Dickson, Mark) mcdpaxil.stanford.edu  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

FEATURES  
 source Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILNI at: http://image.llnl.gov  
 Series: ILNI Plate: 36 Row: n Column: 16.  
 Location/Qualifiers  
 1. 1225

CDs  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:24137 IMAGE:4694856"  
 /issue\_type="Lung"  
 /clone\_lib="NIH\_MGC\_77"  
 /lab\_host="DH10B"  
 /note="Vector: pDNR-LIB"  
 148. 1101  
 /codon\_start=1  
 /product="unknown (protein for MGC:24137)"  
 /protein\_id="AAH2401.1"

/db\_xref="GI:18490240"  
 /translation="MVDPNGSSSATYFFILIGLPGLEAOPFLAPLGLSLYLAVIGN  
 LTIIVYREHSESHLHMYFLCMGLIDLLSTSSPKRLATPENSRTIODADCLIO  
 MPALHLSGMSBESTVLLAMPDPRVAILCHLRLATVILTRTKIGVAATVRGALMAE  
 LPVFLQPFCKSLIISHYCLADQMKACDIRNVNAYGLIIVISALGDSLLISF  
 SYLLIKLVGLIREQAQAFQGVSHVAIVEFVPEFLGDMVRHFRSRRDSPVPI  
 LNIYILVPPVNPPIYGVKTEIKRIRILRPHVATHASEP"  
 BASE COUNT 273 a 331 c 261 g 360 t  
 ORIGIN  
 Query Match 37.8%; Score 1186.4; DB 9; Length 1225;  
 Best Local Similarity 99.5%; Pred. No. 66-252;  
 Matches 1190; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CAGAGAGGCTGTATTTCACTGACAGCTGCGCAGACCTCTCTGAGAGAGACTGACAAAG 60  
 DB 16 CAGAGAGGCTGTATTTCACTGACAGCTGCGCAGACCTCTCTGAGAGAGACTGACAAAG 75  
 QY 61 GGGGTACACATTCCTTCATACAGGTTAGGCTCTACCTGCTGTCTGCTACAGTTC 120  
 DB 76 GGGGTACACATTCCTTCATACAGGTTAGGCTCTACCTGCTGTCTGCTACAGTTC 135  
 QY 121 AGCTTCTCATATAGGTGTGATCCCAATGAGCAATGCAATGCAATGCAATGCAATGCA 180  
 DB 136 AGCTTCTCATATAGGTGTGATCCCAATGAGCAATGCAATGCAATGCAATGCAATGCA 195  
 QY 181 ATGAGCTCTCTGCTTTAGAGAGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 DB 196 ATGAGCTCTCTGCTTTAGAGAGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 255  
 QY 241 TACCTATGCTGTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 256 TACCTATGCTGTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 315  
 QY 301 CTGATGAGCCCATGATATATATTTCTTGTGACGCTTTCAGGATTCATCTCATCTCC 360  
 DB 316 CTGATGAGCCCATGATATATATTTCTTGTGACGCTTTCAGGATTCATCTCATCTCC 375  
 QY 361 ACCATCATGAGCCCAATATGAGGCAATCTCTGCTCAATTCATACATCCATCCAGTTT 420  
 DB 376 ACCATCATGAGCCCAATATGAGGCAATCTCTGCTCAATTCATACATCCATCCAGTTT 435  
 QY 421 GATGCTGTCTGTACAGATTTTTCATGACATCCATCTTATCTGAGATGGAATCCAGTG 480  
 DB 436 GATGCTGTCTGTACAGATTTTTCATGACATCCATCTTATCTGAGATGGAATCCAGTG 495  
 QY 481 CTGCTGCGCATAGGCTTTGACCGCTATGTGAGCATCTGTACCCACTGCGCATCCAC 540  
 DB 496 CTGCTGCGCATAGGCTTTGACCGCTATGTGAGCATCTGTACCCACTGCGCATCCAC 555  
 QY 541 GACTTACGTTGCTCTGTGTACCAAAATGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGT 600  
 DB 556 GACTTACGTTGCTCTGTGTACCAAAATGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGT 615  
 QY 601 CTGATGAGCAACCTTCTCTGTCTTATCAAGCAAGCTGCTGCTGCTGCTGCTGCTGCT 660  
 DB 616 CTGATGAGCAACCTTCTCTGTCTTATCAAGCAAGCTGCTGCTGCTGCTGCTGCTGCT 675  
 QY 661 TCCCATTCCTACTGCTACACCAAGATGTCAAGAGTGTGCTGTGTGTGTGTGTGTGTGT 720  
 DB 676 TCCCATTCCTACTGCTACACCAAGATGTCAAGAGTGTGCTGTGTGTGTGTGTGTGTGT 735  
 QY 721 AATGTGTGTATAGGCTTATGTCTATCATCTCCGCAATGAGCTGTGATCTCATCTCAT 780  
 DB 736 AATGTGTGTATAGGCTTATGTCTATCATCTCCGCAATGAGCTGTGATCTCATCTCAT 795  
 QY 781 TCCCTTCATATCTCTTATCTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840  
 DB 796 TCCCTTCATATCTCTTATCTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 855  
 QY 841 AAGGATTTGGACATGTCGCTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900  
 DB 856 AAGGATTTGGACATGTCGCTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 915

QY	901	ATTGATATGTCATGTGTCATCGCTTTAGCAAGCGCGTGAATCTCCGCTGCCGTATC	960
Db	916	ATTGATATGTCATGTGTCATCGCTTTAGCAAGCGCGTGAATCTCCGCTGCCGTATC	975
QY	961	TTGGCCAAATATCTATCTGTCTGTCTCTCTGTCTCAACCCAAATTGTCTATGAGTGAAG	1020
Db	976	TTGGCCAAATATCTATCTGTCTGTCTCTCTGTCTCAACCCAAATGTCTATGAGTGAAG	1035
QY	1021	ACAAAGAGATTCGACAGCGCATCCTTGACATTTTCCATGTGGCCACACAGGCTTCAAG	1086
Db	1036	ACAAAGAGATTCGACAGCGCATCCTTGACATTTTCCATGTGGCCACACAGGCTTCAAG	1095
QY	1081	CCCTNGGTGTCAGTATCAAACTTCTTTTCCATTCAGAGTCCCTGATTCAGATTTTAA	1140
Db	1096	CCCTNGGTGTCAGTATCAAACTTCTTTTCCATTCAGAGTCCCTGATTCAGATTTTAA	1155
QY	1141	GTTACATTTTGGAAAGACAGTATTCAGAAAAAAATTCCTTAATATAAAATATCA	1196
Db	1156	GTTACATTTTGGAAAGACAGTATTCAGAAAAAAATTCCTTAATATAAAATATCA	1211

RESULT	14		
AX646553			
LOCUS	AX646553	1354 bp	DNA
DEFINITION	Sequence 745 from Patent EP1270724.		linear
ACCESSION	AX646553		
VERSION	AX646553.1	GI:28798936	
KEYWORDS			
SOURCE	Homo sapiens (human)		

ORGANISM  
*Homo sapiens*  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1  
Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.  
Guanosine triphosphate-binding protein coupled receptors  
Patent: EP 1270724-A 745 02-JAN-2003.  
National Institute of Advanced Industrial Science and Technology  
(NII) ; Center for Advanced Science and Technology Incubation, Ltd.  
(JP)

```
FEATURES
  source
    1..1354
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
  CDS
    201..1154
```

```

/note="unnamed protein product"
/codon_start=1
/protein_id="CAD69403.1"
/db_xref="gi:28798937"
/translation="MVDPNGNESSATYFIIILGLDLEAQFWLAFPLCSLYLLAVLIGN
LTIIVIRTEHSLEHPYIFLCMLSGDIIILSTSMKRLAIFENSTIIQFDACILQ
MFAIHSLSGMSSTILAMAFDRVYAIHPLEHAIVTLPRYTKGVAAVGALMLP
LPVILKQLPCRSNII.SHSYCHODWKLACDDI.RANVVG.IYIISAI.GDLSLISE
SYLLIKTVIGLITREAAQAKGCTGSHVCAENIYVYFIFGSMWRHSKRKRSDELPIYI
LANTYLLVPEPLNPIYVGVKTKETKRLRLLRHVATPASEP"

```

	Query Match	37.8%	Score 1186.2;	DB 6;	Length 1354;
	Best Local Similarity	99.3%	Pred. No.6-252;		
	Matches 1202; Conservative	0;	Mismatches 8;	Indels 1;	Gaps 1;
<hr/>					
Qy	77	TCCATACGGTTGAGCCTCTTACCCTCGCTGCTGTGCATCAAGTTCAAGCTTCTTCATGAATGG	136		
Db	145	TCTTTGATTCCAACTCTTACTCTCCGTCGTCTGTACAAGTTAAGCTTCTTCTCATGAATGG	204		
<hr/>					
Qy	137	TGATATCCCAATGAGCAATGTAATCCAGTGCTACATACTTATCATGTAATAGAGCCTCCCTGATT	196		
Dd	205	TGATATCCCAATGAGCAATGTAATCCAGTGCTACATACTTATCATGTAATAGAGCCTCCCTGATT	264		
<hr/>					
Qy	197	TAGAAGAGCTCAATTTCTGTTGGCCTTCCATTGTGTCCCTTACCTTAATGTCTGTGC	256		

Db	285	TAGAAGAGGCTCAGTTCTGTGGTTTGGCCCTCCACTTGTGGTCCCTTCACTTAATTTAGTGTGG	322
Qy	285	TAGAAGAGGCTCAGTTCTGTGGTTTGGCCCTCCACTTGTGGTCCCTTCACTTAATTTAGTGTGG	322
Qy	287	TAGGTAACCTTGACAACTCACTTCACTTATTTGTGGGAGCTAGAGCAACGCTGTGATGAGCCCATGT	316
Db	325	TAGGTAACCTTGACAACTCACTTCACTTATTTGTGGGAGCTAGAGCAACGCTGTGATGAGCCCATGT	384
Qy	317	ATAATATTTCTTTGCACTGCTTTCAGGCACTTGACATCCTCATCTTCCAACTCATCATGCCCA	376
Db	385	ATAATATTTCTTTGCACTGCTTTCAGGCACTTGACATCCTCATCTTCCAACTCATCATGCCCA	444
Qy	377	AAATGCTGGCCATCTTCTGTGTTCAATTCOACTACCATTCAGTTGATGCTTGTCTCTAC	436
Db	445	AAATGCTGGCCATCTTCTGTGTTCAATTCOACTACCATTCAGTTGATGCTTGTCTCTAC	504
Qy	457	AGATTTTGGCACTCCACTCCTTATCTGGAGATGGAATCCACAATGTGCTGGGCAATGGCTT	496
Db	505	AGATTTTGGCACTCCACTCCTTATCTGGAGATGGAATCCACAATGTGCTGGGCAATGGCTT	564
Qy	497	TTGACCGCTATGTGGCCCATCTGTGTCACCCCACTGCGCCATGCCACAGTACTTAAGCTTGGCTC	556
Db	585	TTGACCGCTATGTGGCCCATCTGTGTCACCCCACTGCGCCATGCCACAGTACTTAAGCTTGGCTC	624
Qy	557	GTGTACACCAAAATTTGT	616
Db	625	GTGTACACCAAAATTTGT	684
Qy	617	CTGTCTTCATCAAGAGAGCTGACCTTCTCCGCTCCCAATATCCTTTCCTTCTTACTGAC	676
Db	685	CTGTCTTCATCAAGAGAGCTGACCTTCTCCGCTCCCAATATCCTTTCCTTCTTACTGAC	744
Qy	677	TACACCAAGATGTCTAGAGCTGACCTGTGTATGATATCCGGGTCAATGTCTGTATGACC	736
Db	745	TACACCAAGATGTCTAGAGCTGACCTGTGTATGATATCCGGGTCAATGTCTGTATGACC	804
Qy	727	TTATGCTATATCTTCGCGCAATGGCCCTGGAATCACTCTCATATTCCTCTCATATCTGAC	796
Db	805	TTATGCTATATCTTCGCGCAATGGCCCTGGAATCACTCTCTCATATTCCTCTCTATCTGAC	864
Qy	797	TTATTTCTTAAGACTGTGTGGGCTTGGACAGCTGAAAGCCAGGCCAAAGGATTTGGCACTT	856
Db	865	TTATTTCTTAAGACTGTGTGGGCTTGGACAGCTGAAAGCCAGGCCAAAGGATTTGGCACTT	924
Qy	857	GCGTCTCTCAAGT	916
Db	925	GCGTCTCTCAAGT	984
Qy	917	TGCATTCGCTTGAAGAGGCGGTGACTCTCCGCTGCGCGCTCATCTTGGCCAAATATCTATC	976
Db	985	TGCATTCGCTTGAAGAGGCGGTGACTCTCCGCTGCGCGCTCATCTTGGCCAAATATCTATC	1044
Qy	977	TGCTGTGTTCTCTCTGTGTCTCAACCCAAATGTCTATGTAGGTGAGCAAAAGAGATTTGCAC	1032
Db	1045	TGCTGTGTTCTCTCTGTGTCTCAACCCAAATGTCTATGTAGGTGAGCAAAAGAGATTTGCAC	1104
Qy	1037	AGCGCATCTTCGACTTTTTCATGTGGGCAACACAGCTTCAGAGGCTTAGGTGAGTGA	1096
Db	1105	AGCGCATCTTCGACTTTTTCATGTGGGCAACACAGCTTCAGAGGCTTAGGTGAGTGA	1168
Qy	1097	TCAAACCTTTTTCATTCAGAGTCTCTGTGATTCAGATTTTAAATGTAACTTTTGGAG	1156
Db	1165	TCAAACCTTTTTCATTCAGAGTCTCTGTGATTCAGATTTTAAATGTAACTTTTGGAG	1224
Qy	1157	ACAGTATTCAGAAAAAATTTCTCTTAATAAAAAATCACTCGATCGATCCTTCAATATAGA	1212
Db	1225	ACAGTATTCAGAAAAAATTTCTCTTAAT -AAAAATACACTCGATCCTTCAATATAGA	1284
Qy	1217	AACGTGTGGGAATCTTCATTTTTCATATATATATTTCTTCTTGTGTTTCTGTCTACAT	1276
Db	1284	AACGTGTGGGAATCTTCATTTTTCATATATATATTTCTTCTTGTGTTTCTGTCTACAT	1344
Qy	1277	ATAATTAATAA 1287	
Db	1344	ATAATTAATAA 1354	

RESULT 15  
AB065787 1354 bp DNA linear PRI 23-Jul-2002  
LOCUS Homo sapiens gene for seven transmembrane helix receptor, complete  
DEFINITION cds, isolate:CBRC7TM\_350.  
ACCESSION AB065787  
VERSION AB065787.1 GI:21928839  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Suwa, M., Sato, T., Okouchi, I., Arita, M., Futami, K., Matsumoto, S.,  
Tsutsuni, S., Aburatani, H., Asai, K. and Akiyama, Y.  
TITLE Genome-wide discovery and analysis of human seven transmembrane  
helix receptor genes  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1354)  
AUTHORS Suwa, M.  
TITLE Direct Submission  
JOURNAL Submitted (11-Jul-2001) Makiko Suwa, Computational Biology Research  
Center (CBRC), National Institute of Advanced Industrial Science  
and Technology (AIST), 2-4-16 Aomi Koto-ku, Tokyo 135-0064, Japan  
(E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/,  
Tel:81-3-3599-8080, Fax:81-3-3599-8081)  
COMMENT This sequence is a seven transmembrane helix receptor candidate  
predicted from the whole human genome sequences using our automated  
system (GeneDecoder), sequence search, motif-domain assignment and  
transmembrane helix prediction.  
And the sequence is submitted by the collaborative project between  
[Computational Biology Research Center (CBRC), National Institute  
of Advanced Industrial Science and Technology (AIST)] and [Genome  
Science Division, Research Center for Advanced Science and  
Technology (RCAST), University of Tokyo].  
FEATURES  
source location/Qualifiers  
1..1354  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/isolate="CBRC7TM\_350"  
/db\_xref="taxon:9606"  
/chromosome="11"  
201..1154  
/codon\_start=1  
/evidence=not experimental  
/product="seven transmembrane helix receptor"  
/protein\_id="BAC06006.1"  
/db\_xref="GI:21928840"  
/translation="MVDPNGNSSATYFIILGPGLEAGFWLAFPLCSLYLTVAVGN  
LTIIVYRTEHSHPYIFLCMSGDIILISTSMKMAIFWPSLTIQFACILQ  
MEALHLSGMSVTLAMAFDYALICPHRAIVLPLPRVIGIYVAAGVAGALNAP  
LPVFIKDLPCFRNLSHSYCLHODVKKACDIDRVVVVGLVITISAGLDLISF  
SYLLIKTVLGLTEAAKAKAFGTCVSHKAVCFIYVDFGLSVNHFPSKRSDPLVLI  
LANTYLPVPLNIVYGVATKTKRKLILFPHATNASP"  
BASE COUNT 297 a 338 c 278 g 441 t  
ORIGIN  
Query Match 37.8%, Score 1186.2; DB 9; Length 1354;  
Best Local Similarity 99.3%; Pred. No. 6.6e-252;  
Matches 1202; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
QY 77 TCCTACGCGTGGACCTCTACCTGCGTGGTGTGTACAGCTTTCATGATGG 136  
DB 145 TCTTGATTCACCTCTACCTGCGTGGTGTGTACAGCTTTCATGATGG 204  
QY 137 TGAATCCCATGGAATGATCCAGTGTACATCTCATCTTAATAGGCTCCCTGTT 196  
DB 205 TGAATCCCATGGAATGATCCAGTGTACATCTCATCTTAATAGGCTCCCTGTT 264  
QY 197 TAGAAGAGGCTCACTTCTGCTGCTCCCATCTTCTCATCTTATTTGCTGTGC 256

DB 265 TAGAAGAGGCTCACTTCTGCTGCTCCCATCTTCTCATCTTATTTGCTGTGC 324  
QY 257 TAGTAACCTGACATCATCTCATCTTATTTGCTGCTGCTGCTGCTGCTGCTGCT 316  
DB 325 TAGTAACCTGACATCATCTCATCTTATTTGCTGCTGCTGCTGCTGCTGCTGCT 384  
QY 317 ATATATTTCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 376  
DB 385 ATATATTTCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 444  
QY 377 AAATGCTGCAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436  
DB 445 AAATGCTGCAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 504  
QY 437 AGATTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496  
DB 505 AGATTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564  
QY 497 TTGACCGCTATGTGGCATCTGTCAACCACTGGCCATGCCAGTACTTACGTTGCTC 556  
DB 565 TTGACCGCTATGTGGCATCTGTCAACCACTGGCCATGCCAGTACTTACGTTGCTC 624  
QY 557 GTGTCAACCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616  
DB 625 GTGTCAACCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684  
QY 617 CTGTCTTATCAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676  
DB 685 CTGTCTTATCAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744  
QY 677 TACACCAAGATTCATGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736  
DB 745 TACACCAAGATTCATGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804  
QY 737 TATGCTATCATCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796  
DB 805 TATGCTATCATCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864  
QY 797 TATGCTATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856  
DB 865 TATGCTATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924  
QY 857 GCGTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916  
DB 925 GCGTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984  
QY 917 TGCATCGCTTATGCAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976  
DB 985 TGCATCGCTTATGCAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044  
QY 977 TCGTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036  
DB 1045 TCGTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104  
QY 1037 AGCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096  
DB 1105 AGCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1164  
QY 1097 TCAACTCTCTTTCATTCAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156  
DB 1165 TCAACTCTCTTTCATTCAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1224  
QY 1157 ACAGATTCAGAAAAAATTTCTTAAATAAATAACACTGATCTTCAATATGA 1216  
DB 1225 ACAGATTCAGAAAAAATTTCTTAAATAAATAACACTGATCTTCAATATGA 1283  
QY 1217 AACTGTTGGGAAATCTCATTTTTCATATATTTTCTTCTTCTTCTTCTGCTCAT 1276  
DB 1284 AACTGTTGGGAAATCTCATTTTTCATATATTTTCTTCTTCTTCTTCTGCTCAT 1343  
QY 1277 ATATATTTAA 1287

Tue Feb 10 06:04:23 2004

us-10-017-066a-1.rge

Page 23

Db 1344 ATATTATTAA 1354

Search completed: February 9, 2004, 19:19:58  
Job time : 1169 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 16:17:31 ; Search time 42 Seconds  
(without alignments)  
1198.008 Million cell updates/sec

Title: US-10-017-066A-2

Sequence: 1 MVDPNGENESSATYFIILGLP.....KEIRQRILRLFHVATHASEP 317

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1617	100.0	317	22	AA85002
2	1614	99.8	317	23	ABG61886
3	1614	99.8	318	21	AAV92365
4	1614	99.8	318	22	AAU69951
5	1614	99.8	318	22	ABR44533
6	1614	99.8	318	22	AAU24561
7	1614	99.8	318	22	AAV01306
8	1614	99.8	318	22	AAE02495
9	1614	99.8	318	22	AAU71659

10	1614	99.8	318	23	ABP95674	Human GPCR polypep
11	1614	99.8	318	23	AAU95411	Human p35p protei
12	1614	99.8	318	23	AAU95746	Human olfactory an
13	1614	99.8	318	23	AAU95181	G-coupled olfactor
14	1614	99.8	318	24	ABU71842	Prostate cancer as
15	1614	99.8	318	22	ABP81974	Human GPCR7b polyp
16	1607	99.4	318	22	ABR44534	Human GPCR7c polyp
17	1607	99.4	318	22	ABR44535	Human olfactory re
18	979	60.5	315	22	AAU71762	Human G-protein co
19	979	60.5	318	23	ABG68142	Human G-protein co
20	979	60.5	320	22	AAU99820	Human prostate CDN
21	979	60.5	320	22	AAU24557	Human olfactory re
22	979	60.5	320	22	AAU01175	Human prostate-spe
23	979	60.5	320	22	AAU99060	Human prostate-spe
24	979	60.5	320	22	AAU85135	Human cancer spect
25	979	60.5	320	23	ABP95671	Human GPCR polypep
26	979	60.5	320	23	ABP51576	Human G protein co
27	979	60.5	320	23	ABG61814	Prostate cancer-as
28	979	60.5	320	23	AAE22713	Human prostate sp
29	979	60.5	320	23	ABR95280	Human p790p protei
30	979	60.5	320	23	AAU95745	Human olfactory an
31	979	60.5	320	23	AAU85177	G-coupled olfactor
32	979	60.5	320	24	ABU71711	Prostate cancer sp
33	979	60.5	320	24	ABP81744	Human olfactory re
34	979	60.5	368	22	AAU75532	Human colon cancer
35	977	60.4	320	22	AAU85003	Amino acid sequenc
36	965.5	59.7	320	23	AAE22712	Human prostate sp
37	964.5	59.6	320	18	AAU01730	Human G-protein re
38	964.5	59.6	320	19	AAU56641	G-protein coupled
39	964.5	59.6	320	22	AAU85004	Prostate-specific
40	964.5	59.6	320	23	ABG96334	Human ovarian canc
41	964.5	59.6	320	23	AAU95557	Human olfactory an
42	937.5	58.0	316	23	ABP61149	Human GPCR protei
43	937.5	58.0	324	22	AAU24556	Human olfactory re
44	937.5	58.0	324	23	ABP95673	Human GPCR polypep
45	937.5	58.0	324	23	ABP61135	Human GPCR protei

## ALIGNMENTS

RESULT 1	
ID	AA85002 standard; Protein; 317 AA.
AA85002	
AC	AA85002;
DT	06-AUG-2001 (first entry)
XX	Human G-protein coupled receptor, PHOR-1.
DE	
XX	G-protein-coupled receptor; prostate; cancer; PHOR-1; kidney; uterine;
KM	cervical; stomach; rectal; cytosolic; vaccine; cell function regulator;
KM	human; prostate homologue of olfactory receptor-1.
XX	
OS	Homo sapiens.
XX	
PN	WO200125434-A1.
PD	12-APR-2001.
XX	
PF	05-OCT-2000; 2000WO-US27543.
XX	
PR	05-OCT-1999; 99US-0157902.
XX	
PA	(UROG-) UROGENESYS INC.
XX	
PI	Raitano AB, Afar DEH, Jakobovits A, Farris M, Hubert RS;
XX	Mitchell SC, Safiran DC;
DR	WPI; 2001-367230/38.
XX	
XX	N-PSDB; AAF83880.



PT Novel gene designated PHOR-1, a G-protein-coupled receptor up-regulated  
 PT in prostate cancer, useful as diagnostic marker and therapeutic target  
 PT for cancers of prostate, kidney, uterus  
 PS Claim 1, Fig 1A-D, 139pg; English.  
 XX  
 CC The invention relates to a novel G-protein-coupled receptor up-regulated  
 CC in prostate cancer, termed PHOR-1. The encoding cDNA is contained in  
 CC plasmid designated p10P3A1 deposited with ATCC as Accession No. PTA-312.  
 CC PHOR-1 polypeptides and polynucleotides are useful for diagnosing the  
 CC presence of cancer, especially prostate, kidney, uterine, cervical,  
 CC stomach or rectal cancer by determining and comparing the level of the  
 CC protein or mRNA expression in test and normal tissue samples.  
 CC Pharmaceutical compositions comprising PHOR-1 is useful for treating  
 CC cancer. PHOR-1 proteins are useful for identifying ligands and other  
 CC agents and cellular constituents that binds to PHOR-1 gene product and  
 CC for generating antibodies which are useful in diagnostic, prognostic and  
 CC imaging methodologies and for the treatment of prostate cancer. Cell  
 CC lines expressing PHOR-1 are useful for identifying protein-protein  
 CC interactions mediated by PHOR-1. The present sequence represents the  
 CC human PHOR-1 (prostate homologue of olfactory receptor-1).

XX  
 SQ Sequence 317 AA;

Query Match 100.0%; Score 1617; DB 22; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 2,86-169;  
 Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDPNNGSSATYFLLIGLPGLEAOFWLAFLCSLYLAVGNLTIIYVTEHSLHEP 60  
 DB 1 MVDPNNGSSATYFLLIGLPGLEAOFWLAFLCSLYLAVGNLTIIYVTEHSLHEP 60  
 QY 61 MYFLCMLSGIDILISTSSMPKMLAFWNSSTTIQFDACLQIPAIHSLSGNESTVLLAM 120  
 DB 61 MYFLCMLSGIDILISTSSMPKMLAFWNSSTTIQFDACLQIPAIHSLSGNESTVLLAM 120  
 QY 121 AEDRYVAICHPLRHATVLTLPRTYKIGVAAVVRGAALMAPLVPFIKQLPFCRSNILLSHY 180  
 DB 121 AEDRYVAICHPLRHATVLTLPRTYKIGVAAVVRGAALMAPLVPFIKQLPFCRSNILLSHY 180  
 QY 121 AEDRYVAICHPLRHATVLTLPRTYKIGVAAVVRGAALMAPLVPFIKQLPFCRSNILLSHY 180  
 DB 121 AEDRYVAICHPLRHATVLTLPRTYKIGVAAVVRGAALMAPLVPFIKQLPFCRSNILLSHY 180  
 QY 181 CHODVMKACDDIRVNVVYGLIVTISAIGDLSLISFSYLLIKTVGLTREAOAKAFG 240  
 DB 181 CHODVMKACDDIRVNVVYGLIVTISAIGDLSLISFSYLLIKTVGLTREAOAKAFG 240  
 QY 181 CHODVMKACDDIRVNVVYGLIVTISAIGDLSLISFSYLLIKTVGLTREAOAKAFG 240  
 DB 181 CHODVMKACDDIRVNVVYGLIVTISAIGDLSLISFSYLLIKTVGLTREAOAKAFG 240  
 QY 241 TCVSHVCAVFIYVPIGLSMVHRSKRDSPLPYILANIYLLVPPVLPVIYGVKTKEI 300  
 DB 241 TCVSHVCAVFIYVPIGLSMVHRSKRDSPLPYILANIYLLVPPVLPVIYGVKTKEI 300  
 QY 301 RORILRLPHVATHASEP 317  
 DB 301 RORILRLPHVATHASEP 317

RESULT 2  
 ID ABG61866 standard; Protein; 317 AA.  
 XX  
 AC ABG61866;  
 XX  
 DT 15-AUG-2002 (first entry)  
 XX  
 DE Prostate cancer-associated protein #87.  
 XX  
 KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.  
 XX  
 OS Mammalia.  
 XX  
 PN WO200230268-A2.  
 XX  
 PD 18-APR-2002.  
 XX  
 PF 12-OCT-2001; 2001WO-US32045.  
 XX

PR 13-OCT-2000; 2000US-0687576.  
 PR 08-DEC-2000; 2000US-0733288.  
 PR 08-DEC-2000; 2000US-0733742.  
 PR 24-JAN-2001; 2001US-263957P.  
 PR 16-MAR-2001; 2001US-276791P.  
 PR 16-MAR-2001; 2001US-276888P.  
 PR 06-APR-2001; 2001US-281922P.  
 PR 24-APR-2001; 2001US-286214P.  
 PR 30-APR-2001; 2001US-0847046.  
 PR 04-MAY-2001; 2001US-28589P.  
 XX  
 PA (BOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
 XX  
 DR WPI; 2002-471335/50.  
 DR N-FSDB; ABK92202.  
 XX

PT Detecting a prostate cancer-associated transcript in a cell in a  
 PT patient, useful for diagnosing prostate cancer (PC) or screening  
 PT modulators of PC, by determining if prostate cancer-associated genes  
 PT are expressed in a prostate tissue  
 XX  
 PS Claim 27; Page 373; 436pg; English.

CC The present invention relates to methods of detecting a prostate  
 CC cancer-associated transcript in a cell from a patient. The method  
 CC comprises contacting a biological sample from the patient with  
 CC prostate cancer-associated polynucleotides (designated PC genes) that  
 CC selectively hybridize to a sequence that is at least 80% identical  
 CC to them. The prostate cancer-associated polynucleotide sequences  
 CC are differentially expressed in prostate tumour tissue or in  
 CC prostate cancer and are derived from the tissues of various  
 CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).  
 CC The methods of the invention are useful for diagnosing and treating  
 CC prostate cancer in mammals. The prostate cancer-associated genes are  
 CC useful for diagnosing or treating prostate cancer, as well as for  
 CC identifying modulators of prostate cancer or agents that inhibit  
 CC prostate cancer. The nucleic acid sequences are particularly useful  
 CC in gene therapy, as a vaccine or in antisense applications.  
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins.  
 XX

SQ Sequence 317 AA;

Query Match 99.8%; Score 1614; DB 23; Length 317;  
 Best Local Similarity 99.7%; Pred. No. 5,98-169;  
 Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDPNNGSSATYFLLIGLPGLEAOFWLAFLCSLYLAVGNLTIIYVTEHSLHEP 60  
 DB 1 MVDPNNGSSATYFLLIGLPGLEAOFWLAFLCSLYLAVGNLTIIYVTEHSLHEP 60  
 QY 61 MYFLCMLSGIDILISTSSMPKMLAFWNSSTTIQFDACLQIPAIHSLSGNESTVLLAM 120  
 DB 61 MYFLCMLSGIDILISTSSMPKMLAFWNSSTTIQFDACLQIPAIHSLSGNESTVLLAM 120  
 QY 121 AEDRYVAICHPLRHATVLTLPRTYKIGVAAVVRGAALMAPLVPFIKQLPFCRSNILLSHY 180  
 DB 121 AEDRYVAICHPLRHATVLTLPRTYKIGVAAVVRGAALMAPLVPFIKQLPFCRSNILLSHY 180  
 QY 121 AEDRYVAICHPLRHATVLTLPRTYKIGVAAVVRGAALMAPLVPFIKQLPFCRSNILLSHY 180  
 DB 121 AEDRYVAICHPLRHATVLTLPRTYKIGVAAVVRGAALMAPLVPFIKQLPFCRSNILLSHY 180  
 QY 181 CHODVMKACDDIRVNVVYGLIVTISAIGDLSLISFSYLLIKTVGLTREAOAKAFG 240  
 DB 181 CHODVMKACDDIRVNVVYGLIVTISAIGDLSLISFSYLLIKTVGLTREAOAKAFG 240  
 QY 181 CHODVMKACDDIRVNVVYGLIVTISAIGDLSLISFSYLLIKTVGLTREAOAKAFG 240  
 DB 181 CHODVMKACDDIRVNVVYGLIVTISAIGDLSLISFSYLLIKTVGLTREAOAKAFG 240  
 QY 241 TCVSHVCAVFIYVPIGLSMVHRSKRDSPLPYILANIYLLVPPVLPVIYGVKTKEI 300  
 DB 241 TCVSHVCAVFIYVPIGLSMVHRSKRDSPLPYILANIYLLVPPVLPVIYGVKTKEI 300  
 QY 301 RORILRLPHVATHASEP 317  
 DB 301 RORILRLPHVATHASEP 317

Key	Location/Qualifiers
1..26	
/label= signal_peptide	
8	
/note= "potential glycosylation site"	
34..52	
/label= transmembrane	
45	
/note= "potential glycosylation site"	
57	
/note= "potential phosphorylation site"	
62..83	
/label= signature_sequence	
/note= "G-protein coupled receptor"	
70	
/note= "potential phosphorylation site"	
91	
/note= "potential glycosylation site"	
113..129	
/label= signature_sequence	
/note= "G-protein coupled receptor"	
111	
/note= "potential phosphorylation site"	
180..194	
/label= signature_sequence	
/note= "G-protein coupled receptor"	
202..225	
/label= transmembrane	
240..255	
/label= signature_sequence	
/note= "G-protein coupled receptor"	
267	
/note= "potential phosphorylation site"	
272	
/note= "potential phosphorylation site"	

MO200020590-A2.

13-APR-2000.

06-OCT-1999; 99WO-US23317.

06-OCT-1998; 98US-0167219.

06-OCT-1998; 98US-0172211.

11-MAY-1999; 99US-0133585.

(INCYTE PHARM INC.

Tang YT, Yue H, Lai P, Bandhan O, Au-oung J, Reddy R, Corley NC, Giesler KJ, Gorgone GA, Baughn MR, Azimzai Y, WPI: 2000-338934/28.

N-PSDB; AAA09351.

Novel human G-protein coupled receptor proteins used in the diagnosis,

converted to protein

from signal / Ab seq.

PT	treatment and prevention of nervous system disorders,
FT	autoimmune/inflammatory disorders, and cell proliferative disorders
XX	such as cancer
PS	Claim 1; Page 71-72; 84pp; English.
XX	
CC	This sequence encodes human G-protein coupled receptor protein (GCRP) 5.
CC	The GCRP polypeptides, polynucleotides, antibodies, antagonists and
CC	agonists may be administered to human patients for the diagnosis,
CC	treatment and prevention of nervous system disorders (e.g. epilepsy,
CC	stroke, neoplasms, Alzheimer's disease), autoimmune or
CC	inflammatory disorders, complications of cancer, hemodialysis and
CC	extracorporeal circulation, and cell proliferative disorders. They are
CC	also used to treat or prevent disorders associated with decreased or
CC	increased expression or activity of GCRP.
XX	
SQ	Sequence    318 AA; <i>5'cag ATC</i>
Query Match	99.8%; Score 1614; DB 21; Length 318;
Best Local Similarity	99.7%; Pred. No. 5,9e-169;
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0	
QY	1 MDVDPNGNESSATYFILLIGLPGLEEAQFWLAAPLCGLYLIAVLGNLTIIIVRTSHSLHEP 60
DB	2 MDVPNGNESSATYFILLIGLPGLEEAQFWLAAPLCGLYLIAVLGNLTIIIVRTSHSLHEP 61
QY	61 MYTFLCNLSGIDILISTSMRPKMLIFPNSTTIOFACLIQTAHISLSMESTVLLAM 120
DB	62 MYTFLCNLSGIDILISTSMRPMQLIFPNSTTTQFDACLQMFAHISLSMESTVLLAM 121
QY	121 AEDRYVAICHPLRHATVTLTPRVTKIGVAAVVRGAALMAPLPVPIKQLPCFSNLIASHSY 180
DB	122 AEDRYVAICHPLRHATVTLTPRVTKIGVAAVVRGAALMAPLPVPIKQLPCFSNLIASHSY 181
QY	181 CLHQDYMKLACDDIRNNVNYGIVITSAIGDSLISFSYLLIKTVLGITREAQAAG 240
DB	182 CLHQDYMKLACDDIRNNVNYGIVITSAIGDSLISFSYLLIKTVLGITREAQAAG 241
QY	241 TCVSHCAVFIEYVFFIGISWVHRSKRDSPPLVIANIYLVPPVLNPVIYGVKTKEI 300
DB	242 TCVSHCANFIIEYVFFIGISWVHRSKRDSPPLVIANIYLVPPVLNPVIYGVKTKEI 301
QY	301 RQRILRFVATHASEP 317
DB	302 RQRILRFVATHASEP 318
RESULT 4	
ID	AU69951
TD	AU69951 standard; Protein; 318 AA.
XX	
AC	AAU69951;
DT	30-JAN-2002 (first entry)
XX	
DE	Human prostate cDNA encoded protein #85.
KX	Human; prostate cancer; cytosolic; immunostimulant; tumour; immunogen.
XX	
OS	Homo sapiens.
XX	
PB	MO200173032-A2.
PD	04-OCT-2001.
PF	27-MAR-2001; 2001WO-US09919.
PR	27-MAR-2000; 2000US-0536857.
PR	09-MAY-2000; 2000US-0568100.
PR	12-MAY-2000; 2000US-0570337.
PR	13-JUN-2000; 2000US-0593793.
PR	27-JUN-2000; 2000US-0605783.
PR	10-AUG-2000; 2000US-0636215.



CC control samples. They are useful to identify agents binding polypeptide  
 CC (e.g. cellular receptors or downstream effectors) and/or agents  
 CC modulating cellular polypeptide expression or activity, useful as  
 CC antagonists and agonists in disease treatment.

XX Sequence 318 AA;

Query Match 99.8%; Score 1614; DB 22; Length 318;  
 Best Local Similarity 99.7%; Pred. No. 5.9e-169;  
 Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDNNGESSATYFLLIGLPGLERAOFWLAFPLCSLYLAVLGNLTIIYVREHSLHEP 60  
 DB 2 MVDNNGESSATYFLLIGLPGLERAOFWLAFPLCSLYLAVLGNLTIIYVREHSLHEP 61  
 QY 61 MYIFLCMSGIDILISTSMRPMALIFWENSTTIQFACLLQIFAHISLSGMESTVLLAM 120  
 DB 62 MYIFLCMSGIDILISTSMRPMALIFWENSTTIQFACLLQIFAHISLSGMESTVLLAM 121  
 QY 121 AFDRYVAICHPRLRATVTLTPRVTKIGVAAVVRGALMAPLPVFIKQLPFCRSNILLSHY 180  
 DB 122 AFDRYVAICHPRLRATVTLTPRVTKIGVAAVVRGALMAPLPVFIKQLPFCRSNILLSHY 181  
 QY 181 CLHODVMKACDDIRVNVVYGLIYISAGLDSLLISFSYLLIKTYLGLTREAOAKAFG 240  
 DB 182 CLHODVMKACDDIRVNVVYGLIYISAGLDSLLISFSYLLIKTYLGLTREAOAKAFG 241  
 QY 241 TCVSHVCAVFIFYVPFGLSMVHRFSKRDSPLPVILANITYLLVPVLPNIYGVGVTKEI 300  
 DB 242 TCVSHVCAVFIFYVPFGLSMVHRFSKRDSPLPVILANITYLLVPVLPNIYGVGVTKEI 301  
 QY 301 RORILRLFHVATHASEP 317  
 DB 302 RORILRLFHVATHASEP 318

RESULT 6  
 AAU24561  
 ID AAU24561 standard; Protein; 318 AA.

AC AAU24561;  
 DT 18-DEC-2001 (first entry)  
 DE Human olfactory receptor AOLF48.  
 XX

KM Human olfactory receptor; G protein-coupled receptor; GPCR; odorant;  
 KM food additive; cosmetic; fragrance; pharmaceutical additive.

XX Homo sapiens.

OS WO200168805-A2.

XX PD 20-SEP-2001.

XX PF 13-MAR-2001; 2001WO-US07771.

XX PR 13-MAR-2000; 2000US-0188914.

XX PR 24-MAR-2000; 2000US-0192033.

XX PR 12-APR-2000; 2000US-0198474.

XX PR 24-APR-2000; 2000US-0199335.

XX PR 26-MAY-2000; 2000US-0207702.

XX PR 23-JUN-2000; 2000US-0213849.

XX PR 16-AUG-2000; 2000US-0226534.

XX PR 07-SEP-2000; 2000US-0230732.

XX PR 07-FEB-2001; 2001US-0266662.

XX PA (SENSO-) SENOMIX INC.

XX Zozulya S;

XX WPI; 2001-570867/64.

XX DR N-PSDB; AAS42254.

XX Nucleic acids encoding human olfactory G protein-coupled receptors.  
 PT useful for screening for compounds involved in olfactory sensation,  
 PT where the compounds can be used in the food, pharmaceutical and  
 PT cosmetic industries to customise odours -

XX Claim 60; Page 103-104; 319pp; English.

XX The invention relates to nucleic acids encoding human olfactory  
 CC receptors, OR, (a G protein-coupled receptor, GPCR). The ORs  
 CC specifically recognise molecules, odorants, that elicit specific  
 CC olfactory sensation. The human olfactory receptors and polynucleotides  
 CC encoding them are useful for screening a library of chemical compounds  
 CC for compounds that are involved in olfactory sensation. Modulators of  
 CC their activity are useful for pharmacological and genetic modulation of  
 CC olfactory signalling pathways. Therefore, they can be used in the food,  
 CC pharmaceutical and cosmetic industries to customise odours and  
 CC fragrances. The present sequence is a human olfactory receptor of the  
 CC invention.

XX Sequence 318 AA;

Query Match 99.8%; Score 1614; DB 22; Length 318;  
 Best Local Similarity 99.7%; Pred. No. 5.9e-169;  
 Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDNNGESSATYFLLIGLPGLERAOFWLAFPLCSLYLAVLGNLTIIYVREHSLHEP 60  
 DB 2 MVDNNGESSATYFLLIGLPGLERAOFWLAFPLCSLYLAVLGNLTIIYVREHSLHEP 61  
 QY 61 MYIFLCMSGIDILISTSMRPMALIFWENSTTIQFACLLQIFAHISLSGMESTVLLAM 120  
 DB 62 MYIFLCMSGIDILISTSMRPMALIFWENSTTIQFACLLQIFAHISLSGMESTVLLAM 121  
 QY 121 AFDRYVAICHPRLRATVTLTPRVTKIGVAAVVRGALMAPLPVFIKQLPFCRSNILLSHY 180  
 DB 122 AFDRYVAICHPRLRATVTLTPRVTKIGVAAVVRGALMAPLPVFIKQLPFCRSNILLSHY 181  
 QY 181 CLHODVMKACDDIRVNVVYGLIYISAGLDSLLISFSYLLIKTYLGLTREAOAKAFG 240  
 DB 182 CLHODVMKACDDIRVNVVYGLIYISAGLDSLLISFSYLLIKTYLGLTREAOAKAFG 241  
 QY 241 TCVSHVCAVFIFYVPFGLSMVHRFSKRDSPLPVILANITYLLVPVLPNIYGVGVTKEI 300  
 DB 242 TCVSHVCAVFIFYVPFGLSMVHRFSKRDSPLPVILANITYLLVPVLPNIYGVGVTKEI 301  
 QY 301 RORILRLFHVATHASEP 317  
 DB 302 RORILRLFHVATHASEP 318

RESULT 7

AAU01306  
 ID AAU01306 standard; Protein; 318 AA.

AC AAU01306;

DT 04-OCT-2001 (first entry)

DE P835P amino acid sequence.

KM Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
 KM cyclostatic; gene therapy; metastasis.

XX Homo sapiens.

OS WO200151633-A2.

XX PD 19-JUL-2001.

XX PF 16-JAN-2001; 2001WO-US01574.

XX PR 14-JAN-2000; 2000US-0483672.



CC their corresponding DNA molecules. GPCR is also referred as seven  
 CC transmembrane receptor. G protein-coupled receptor protein is useful for  
 CC treating neurological disorder, particularly schizophrenia. GPCR protein  
 CC is also useful for identifying compounds useful for treating  
 CC schizophrenia. These compounds are also useful for treating other  
 CC neurological and psychiatric diseases, e.g. depression, anxiety, bipolar  
 CC disease, affective disorders, attention deficit hyperactivity disorder/  
 CC attention deficit disorder, epilepsy, neuritis, neuroasthenia, neuropathy,  
 CC neurosis, Alzheimer's disease, Parkinson's disease, migraine and senile  
 CC dementia. The invention also provides genetic screening procedures that  
 CC entail analysing a person's genome with respect to GPCR. The vectors are  
 CC useful for the recombinant production of the GPCR's. The present sequence  
 CC is human CON198 G protein-coupled receptor (GPCR) protein.

XX Sequence 318 AA;

Query Match 99.8%; Score 1614; DB 22; Length 318;

Best Local Similarity 99.7%; Pred. No. 5.9e-169; Mismatches 0; Gaps 0;

Matches 316; Conservative 1; Indels 0; Gaps 0;

QY 1 MVDNNGESSATYFILLGLPGLERAOFWLAFPLCSLYLAVLGNLTIIYVTEHSLHEP 60

DB 2 MVDNNGESSATYFILLGLPGLERAOFWLAFPLCSLYLAVLGNLTIIYVTEHSLHEP 61

QY 61 MYFLCMLSGDILISTSSMPKMLAFWENSTTIQPDACILQIFAHISLGMESTVLLAM 120

DB 62 MYFLCMLSGDILISTSSMPKMLAFWENSTTIQPDACILQIFAHISLGMESTVLLAM 121

QY 121 AFDRYVAICHPLRHATVLTLPRTKIGVAAVVRGAALMAPLPVFIKQLPFCRSNLSHSY 180

DB 122 AFDRYVAICHPLRHATVLTLPRTKIGVAAVVRGAALMAPLPVFIKQLPFCRSNLSHSY 181

QY 181 CLHODVWKLACDDIRVNVVYGLIVISAGLDSLLISFSYLLIKTVLGLTREAOAKAFG 240

DB 182 CLHODVWKLACDDIRVNVVYGLIVISAGLDSLLISFSYLLIKTVLGLTREAOAKAFG 241

QY 241 TCVSHVCAVFIFFYVPIGLSMVHRFSKRDSPLPVILLANTYLLVPPVLPVYGVTKKEI 300

DB 242 TCVSHVCAVFIFFYVPIGLSMVHRFSKRDSPLPVILLANTYLLVPPVLPVYGVTKKEI 301

QY 301 RQRIILRFHVATHASEP 317

DB 302 RQRIILRFHVATHASEP 318

RESULT 9

AAG71659

ID AAG71659 standard; Protein; 318 AA.

AC AAG71659;

DT 30-JUL-2001 (first entry)

DE Human olfactory receptor polypeptide, SEQ ID NO: 1340.

XX Human; olfactory receptor; OR; primary scent determination;

KM secondary scent determination; polypeptide library; odour receptor;

KW scent profile; scent fingerprint; scent representation.

OS Homo sapiens.

XX WO200127158-A2.

XX 19-APR-2001.

PF 06-OCT-2000; 2000MO-US27582.

PR 08-OCT-1999; 99US-0158615.

FR 24-FEB-2000; 2000US-0184809.

PA (DIGI-) DIGISCENTS.

XX (YEDA) YEDA RES & DEV CO LTD.

PI Bellenson J, Smith D, Lancel D, Glusman G, Fuchs T, Yanai I;  
 XX WPI: 2001-290713/30.  
 DR  
 XX  
 PT New polynucleotides which encode polypeptides involved in olfactory  
 PT sensation for identifying olfactory agonists and antagonists -  
 PS Claim 11; Page 826-827; 1857pp; English.

CC The present sequence is an olfactory receptor which is encoded by  
 CC one of a number of novel polynucleotides. The polynucleotides can be  
 CC used in screening for olfactory agonists and antagonists. The methods  
 CC allow for the determination of primary scents and the identification  
 CC of the odour receptors used to detect these primary scents. The methods  
 CC also enable determination of secondary scents and the identification of  
 CC combinations of odour receptors that are involved in detecting such  
 CC secondary scents. This enables the construction of a scent representation  
 CC (also called a scent fingerprint or scent profile), which may be used to  
 CC re-create and edit scents. Libraries of olfactory receptors are useful  
 CC for determining the interaction pattern of a composition with the  
 CC receptors, and can be used for determining differences in the olfactory  
 CC facilities of different individuals.

XX Sequence 318 AA;

Query Match 99.8%; Score 1614; DB 22; Length 318;

Best Local Similarity 99.7%; Pred. No. 5.9e-169; Mismatches 0; Gaps 0;

Matches 316; Conservative 1; Indels 0; Gaps 0;

QY 1 MVDNNGESSATYFILLGLPGLERAOFWLAFPLCSLYLAVLGNLTIIYVTEHSLHEP 60

DB 2 MVDNNGESSATYFILLGLPGLERAOFWLAFPLCSLYLAVLGNLTIIYVTEHSLHEP 61

QY 61 MYFLCMLSGDILISTSSMPKMLAFWENSTTIQPDACILQIFAHISLGMESTVLLAM 120

DB 62 MYFLCMLSGDILISTSSMPKMLAFWENSTTIQPDACILQIFAHISLGMESTVLLAM 121

QY 121 AFDRYVAICHPLRHATVLTLPRTKIGVAAVVRGAALMAPLPVFIKQLPFCRSNLSHSY 180

DB 122 AFDRYVAICHPLRHATVLTLPRTKIGVAAVVRGAALMAPLPVFIKQLPFCRSNLSHSY 181

QY 181 CLHODVWKLACDDIRVNVVYGLIVISAGLDSLLISFSYLLIKTVLGLTREAOAKAFG 240

DB 182 CLHODVWKLACDDIRVNVVYGLIVISAGLDSLLISFSYLLIKTVLGLTREAOAKAFG 241

QY 241 TCVSHVCAVFIFFYVPIGLSMVHRFSKRDSPLPVILLANTYLLVPPVLPVYGVTKKEI 300

DB 242 TCVSHVCAVFIFFYVPIGLSMVHRFSKRDSPLPVILLANTYLLVPPVLPVYGVTKKEI 301

QY 301 RQRIILRFHVATHASEP 317

DB 302 RQRIILRFHVATHASEP 318

RESULT 10

ABP95674

ID ABP95674 standard; Protein; 318 AA.

AC ABP95674;

DT 06-MAR-2003 (first entry)

DE Human GPCR polypeptide SEQ ID NO 158.

XX Human; GPCR; G protein coupled receptor; signal transduction; olfactory;

KM drug development; gustatory; taste; fragrance; receptor.

OS Homo sapiens.

XX WO200216548-A2.

XX 28-FEB-2002.

PF 30-JUL-2001; 2001WO-IB01446.  
 XX 04-AUG-2000; 2000JP-0237818.  
 PR 13-FEB-2001; 2001JP-0034434.  
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
 XX Haga T, Takeda S, Mitaku S;  
 FI WPI; 2002-304118/34.  
 DR N-PEDB; AB242948.  
 XX  
 PT Database global search for G protein-coupled receptors, proteins and  
 PT encoded genes for studying in vivo signal transduction mechanism and  
 PT identifying targets for drug development  
 XX  
 PS Claim 10; SEQ ID NO 158; 97pp + Sequence Listing; Japanese.  
 CC The invention relates to a method for screening G protein-coupled  
 CC receptor (GPCR) genes (AB242870-AB243216) and/or GPCR proteins  
 CC (AB295596-AB295942) by extracting open-reading frames containing 6-8  
 CC transmembrane domains with 250-1000 amino acid residues to give a gene  
 CC homologous with a known GPCR gene. The receptor proteins and encoded  
 CC genes are useful for studying in vivo signal transduction mechanism and  
 CC identifying targets for drug development e.g. based on olfactory and  
 CC gustatory receptors in form of agonists and antagonists by screening  
 CC intrinsic and extrinsic ligands as bitter taste inhibitors, taste  
 CC enhancers and fragrance improvers.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 318 AA;  
 Query Match 99.8%; Score 1614; DB 23; Length 318;  
 Best Local Similarity 99.7%; Pred. No. 5.9e-169;  
 Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDNNGNSATYFLIGLPGLEAOPFLAPPLSLYLIVGLTIIYVREHSHSEP 60  
 DB 2 MVDNNGNSATYFLIGLPGLEAOPFLAPPLSLYLIVGLTIIYVREHSHSEP 61  
 QY 61 MYIFLCMLSGIDILISTSSMPKMLAIWFNSTTIQDPACLLQIPAHISLGMESTVLLAM 120  
 DB 62 MYIFLCMLSGIDILISTSSMPKMLAIWFNSTTIQDPACLLQIPAHISLGMESTVLLAM 121  
 QY 121 AFDRAVAICHPLEAHATVLTLPRTKIGVAAVVGAAIMAPLPVPIKOLPCRSNIISHSY 180  
 DB 122 AFDRAVAICHPLEAHATVLTLPRTKIGVAAVVGAAIMAPLPVPIKOLPCRSNIISHSY 181  
 QY 181 CLHODVMKLACDDIRVNVVYGLIVISAIGLSLLISFSYLLIKTVLGLTREAOAKAFG 240  
 DB 182 CLHODVMKLACDDIRVNVVYGLIVISAIGLSLLISFSYLLIKTVLGLTREAOAKAFG 241  
 QY 241 TCVSHVCAVFLFYVPIGLSMVHRFSKRDSPVLIILANTYLVLPPLNTIYGVTKET 300  
 DB 242 TCVSHVCAVFLFYVPIGLSMVHRFSKRDSPVLIILANTYLVLPPLNTIYGVTKET 301  
 QY 301 RORILRLFHVATHASEP 317  
 DB 302 RORILRLFHVATHASEP 318  
 RESULT 11  
 ABB95411  
 ID ABB95411 standard; Protein; 318 AA.  
 XX AC ABB95411;  
 XX DT 19-JUL-2002 (first entry)  
 DE Human P835P protein SEQ ID NO 920.  
 XX

KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;  
 KW gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002022248-A1.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 12-JAN-2001; 2001US-0759143.  
 XX  
 PR 25-FEB-1997; 97US-0806099.  
 PR 01-AUG-1997; 97US-0904804.  
 PR 09-FEB-1998; 98US-0020956.  
 PR 25-FEB-1998; 98US-0030607.  
 PR 14-JUL-1998; 98US-0115453.  
 PR 23-SEP-1998; 98US-0159812.  
 PR 15-JAN-1999; 99US-0232149.  
 PR 09-APR-1999; 99US-0288946.  
 PR 13-JUL-1999; 99US-0352616.  
 PR 18-NOV-1999; 99US-0439313.  
 PR 14-JAN-2000; 2000US-0483672.  
 PR 27-MAR-2000; 2000US-0536857.  
 PR 09-MAY-2000; 2000US-0568100.  
 PR 12-MAY-2000; 2000US-0570737.  
 PR 13-JUN-2000; 2000US-0593793.  
 PR 27-JUN-2000; 2000US-0605783.  
 PR 10-AUG-2000; 2000US-0636215.  
 PR 29-AUG-2000; 2000US-0651236.  
 PR 06-SEP-2000; 2000US-0657279.  
 PR 02-OCT-2000; 2000US-0679426.  
 PR 10-OCT-2000; 2000US-0685166.  
 XX  
 (XUJ/) XU J.  
 PA (DILL/) DILLON D C.  
 PA (MITC/) MITCHAM J L.  
 PA (HARL/) HARLOCKER S L.  
 PA (JIAN/) JIANG Y.  
 PA (KALO/) KALOS M D.  
 PA (FANG/) FANGER G R.  
 PA (RETT/) RETTER M W.  
 PA (STOL/) STOLK J A.  
 PA (DAYC/) DAY C H.  
 PA (VEDV/) VEDVICK T S.  
 PA (CART/) CARTER D.  
 PA (LISK/) LI S X.  
 PA (WANG/) WANG A.  
 PA (SKET/) SKETKY Y A W.  
 PA (HEPL/) HEPLER W T.  
 PA (HEND/) HENDERSON R A.  
 XX  
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Sketky YAW, Hepler WT, Henderson RA;  
 XX  
 DR WPI; 2002-255649/30.  
 XX  
 PT New prostate-specific polynucleotides for diagnosing and treating  
 PT diseases, in particular prostate cancer, and as markers for the  
 PT progression of cancer  
 XX  
 PS Claim 2; SEQ ID NO 920; 87pp; English.  
 XX  
 CC The present invention provides prostate-specific coding sequences and  
 CC their encoded proteins. These can be used in the diagnosis and treatment  
 CC of cancers, particularly prostate cancer. The present sequence is a  
 CC protein described in the invention.  
 XX  
 SQ Sequence 318 AA;  
 Query Match 99.8%; Score 1614; DB 23; Length 318;  
 Best Local Similarity 99.7%; Pred. No. 5.9e-169;

Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDPNGNSSATYFLLIGLPGLEBAQFWLAFPLCSLYLIANVGNLTIIYVTEHSHHP 60  
 DB 2 MYDPNGNSSATYFLLIGLPGLEBAQFWLAFPLCSLYLIANVGNLTIIYVTEHSHHP 61

QY 61 MYIFLCMLSGIDILISTSMRPMKLAIFWNSSTIIQFDACLIQIFAHSLSGNESTVLLAM 120  
 DB 62 MYIFLCMLSGIDILISTSMRPMKLAIFWNSSTIIQFDACLIQIFAHSLSGNESTVLLAM 121

QY 121 AFDRYVAICHPRLHATVLTLPRTVKIGVAAVRGAALMAPLPVFIKOLPFCRSNLSHSY 180  
 DB 122 AFDRYVAICHPRLHATVLTLPRTVKIGVAAVRGAALMAPLPVFIKOLPFCRSNLSHSY 181

QY 131 CHQDVWKLACDDIRVNVVYGLIYISAIGDLSLISFSYLLIKTYLGLTREQAQAFG 240  
 DB 182 CHQDVWKLACDDIRVNVVYGLIYISAIGDLSLISFSYLLIKTYLGLTREQAQAFG 241

QY 241 TCVSHVCAVFIFYPYFGLSMVHRFSKRDSPLPVILANIYLVPPVLPNPIYGVKTKXI 300  
 DB 242 TCVSHVCAVFIFYPYFGLSMVHRFSKRDSPLPVILANIYLVPPVLPNPIYGVKTKXI 301

QY 301 RORILRLPHVATHASEP 317  
 DB 302 RORILRLPHVATHASEP 318

RESULT 12  
 AAU95746 standard; Protein; 318 AA.

AC AAU95746;  
 XX  
 DT 02-JUN-2002 (first entry)

Human olfactory and pheromone G protein-coupled receptor #33.

Human; olfactory and pheromone G protein coupled; receptor;  
 KM GPCR; transducer; antidepressant; neuroleptic; endocrine; anabolic;  
 KM anorectic; taste; fragrance; food additive; cosmetic; cell migration;  
 KM sterility; psychotic disorder; neurological disorder; anxiety;  
 KM schizophrenia; manic depression; depression; axonal growth;  
 KM menstrual cycle; appetite sexual motivation; sexual attraction;  
 KM aggression.

OS Homo sapiens.  
 XX  
 PN WO200224726-A2.  
 XX  
 PD 28-MAR-2002.  
 XX  
 PF 21-SEP-2001; 2001WO-BE00162.  
 XX  
 PR 22-SEP-2000; 2000EP-0870211.

(CHEM-) CHEMCOM SA.  
 PA  
 PI Velthen A;  
 XX  
 DR MPI: 2002-330013/36.  
 DR N-PSDB; A868633.  
 XX  
 PT Novel pheromone G-protein coupled receptor and receptor-derived  
 PT agonists, antagonists or inhibitors useful in food or cosmetic products  
 PT or in the treatment or prevention of neurological disorders such as  
 PT anxiety and schizophrenia -  
 XX  
 PS Disclosure; Page 710-711; 833pp; English.

The invention relates to olfactory and Pheromone G-protein coupled  
 CC receptor (GPCR) or a protein 95% identical to the GPCR, a specific active  
 CC portion and its encoding polynucleotide. Also included are an agonist,  
 CC antagonist or inhibitor of the GPCR or the polynucleotide, a vector

CC comprising the polynucleotide, a cell transformed by the vector, a non-  
 CC human mammal comprising a partial or total deletion of the polynucleotide  
 CC encoding the receptor and screening (detection and possibly, recovering)  
 CC of compounds which are known or not known to be agonist, antagonists or  
 CC inhibitors of natural compounds to the GPCR. The receptor-derived  
 CC agonists, antagonists, inhibitors or compounds are used as an  
 CC improvement, elimination or substitution of an existing taste and/or a  
 CC fragrance of (or in) the food and/or cosmetic products. They can also be  
 CC used in the preparation of medication in the treatment and/or prevention  
 CC of a mammalian disorder, such as cell migration, sterility, psychotic and  
 CC neurological disorders, including anxiety, schizophrenia, manic  
 CC depression, depression, for promoting axonal growth, nerve cell  
 CC connection and nerve regeneration for modulating male and female  
 CC endocrine functions, hormone production and the menstrual cycle, for the  
 CC prevention or the treatment by stimulation of several mammalian  
 CC behaviours, such as stimulation or suppression of appetite, sexual  
 CC motivation, sexual attraction, aggression and for promoting or  
 CC suppressing chemical communication between organisms. The present  
 CC sequence is a human olfactory and pheromone GPCR protein sequence.

XX Sequence 318 AA;

QY 1 MYDPNGNSSATYFLLIGLPGLEBAQFWLAFPLCSLYLIANVGNLTIIYVTEHSHHP 60  
 DB 2 MYDPNGNSSATYFLLIGLPGLEBAQFWLAFPLCSLYLIANVGNLTIIYVTEHSHHP 61

QY 61 MYIFLCMLSGIDILISTSMRPMKLAIFWNSSTIIQFDACLIQIFAHSLSGNESTVLLAM 120  
 DB 62 MYIFLCMLSGIDILISTSMRPMKLAIFWNSSTIIQFDACLIQIFAHSLSGNESTVLLAM 121

QY 121 AFDRYVAICHPRLHATVLTLPRTVKIGVAAVRGAALMAPLPVFIKOLPFCRSNLSHSY 180  
 DB 122 AFDRYVAICHPRLHATVLTLPRTVKIGVAAVRGAALMAPLPVFIKOLPFCRSNLSHSY 181

QY 181 CHQDVWKLACDDIRVNVVYGLIYISAIGDLSLISFSYLLIKTYLGLTREQAQAFG 240  
 DB 182 CHQDVWKLACDDIRVNVVYGLIYISAIGDLSLISFSYLLIKTYLGLTREQAQAFG 241

QY 241 TCVSHVCAVFIFYPYFGLSMVHRFSKRDSPLPVILANIYLVPPVLPNPIYGVKTKXI 300  
 DB 242 TCVSHVCAVFIFYPYFGLSMVHRFSKRDSPLPVILANIYLVPPVLPNPIYGVKTKXI 301

QY 301 RORILRLPHVATHASEP 317  
 DB 302 RORILRLPHVATHASEP 318

RESULT 13  
 AAU85181 standard; Protein; 318 AA.

AC AAU85181;  
 XX  
 DT 08-MAY-2002 (first entry)

G-coupled olfactory receptor #42.

Human; olfactory G-coupled receptor; sensory perception of odourant;  
 KM odour composition; taste composition.  
 KM  
 XX Homo sapiens.  
 XX  
 PN WO200198526-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PF 22-JUN-2001; 2001WO-US20122.  
 XX  
 PR 22-JUN-2000; 2000US-213812P.



PR 13-MAR-2001; 2001US-0804291.  
 XX (SENO-) SENOMYX INC.  
 XX Zozulya S, Stryer L;  
 XX WPI: 2002-083330/11.  
 DR N-PSDB; ABK37340.  
 PT Representing sensory perception of one or more odourants for the  
 PT identification and design of tastes and odours comprises providing a  
 PT representative group of n olfactory receptors -  
 XX  
 PS Claim 1; Page 77; 182pp; English.  
 CC The invention relates to a method of representing sensory perception of  
 CC one or more odourants. The method comprises: (a) providing a  
 CC representative class of n olfactory receptors or ligand binding domains  
 CC (LBD) of these receptors; (b) measuring values Xi to Xn representative  
 CC of at least one activity of one or more odourants selected from:  
 CC (i) binding one or more odourants to the LBD of at least one of the n  
 CC olfactory receptors; (ii) activating at least one of the n  
 CC olfactory receptors with the one or more odourants; and (iii) blocking at  
 CC least one of the n olfactory receptors with the one or more odourants;  
 CC and (c) generating a representation of sensory perception from the values  
 CC Xi to Xn. The representation of the sensory perception of odourants is  
 CC useful for the design and formulation of odour and taste compositions.  
 CC AU85140-AU85393 represent human olfactory G-coupled receptor amino acid  
 CC sequences of the invention.  
 XX  
 SQ Sequence 318 AA;  
 Query Match 99.8%; Score 1614; DB 23; Length 318;  
 Best Local Similarity 99.7%; Pred. No. 5.9e-169;  
 Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDPNGNESATYFILIGLPGLEAQPWLAFPLCSLYLAVAGNTLIYVTEHSLHP 60  
 DB 2 MDPNGNESATYFILIGLPGLEAQPWLAFPLCSLYLAVAGNTLIYVTEHSLHP 61  
 QY 61 MYFLCMLSGIDLITSSMPKMLAFWNSTTIQPDACLLQIPAIHSLGMSVTLIAM 120  
 DB 62 MYFLCMLSGIDLITSSMPKMLAFWNSTTIQPDACLLQIPAIHSLGMSVTLIAM 121  
 QY 121 AFDRYVALCHPLRHATVLTLPRTKIGVAAVRGALMLAPVFPKQLPFCRSNTLSHY 180  
 DB 122 AFDRYVALCHPLRHATVLTLPRTKIGVAAVRGALMLAPVFPKQLPFCRSNTLSHY 181  
 QY 181 CLHQDVWKLACDDIRNVVYGLIVISAIGDLSLISFSLIILKTVALGLTREAAKAPG 240  
 DB 182 CLHQDVWKLACDDIRNVVYGLIVISAIGDLSLISFSLIILKTVALGLTREAAKAPG 241  
 QY 241 TCVSHCAVEIFVFPVIGLSMWRFSKRDSPPLVILANTYLVLPVNPITYGVKTEI 300  
 DB 242 TCVSHCAVEIFVFPVIGLSMWRFSKRDSPPLVILANTYLVLPVNPITYGVKTEI 301  
 QY 301 RQRILRLFVAVTHASEP 317  
 DB 302 RQRILRLFVAVTHASEP 318  
 RESULT 14  
 ID ABU71842 standard; Protein; 318 AA.  
 AC ABU71842;  
 XX  
 DT 10-JUN-2003 (first entry)  
 XX  
 DE Prostate cancer associated protein #66.  
 XX Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 KW immunogen; cancer; prostate specific antigen; PSA;

KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;  
 KW PSMA.  
 XX Homo sapiens.  
 XX US2002192763-A1.  
 XX 19-DEC-2002.  
 XX 29-JUN-2001; 2001US-0895793.  
 XX 17-APR-2000; 2000US-157455P.  
 XX 04-OCT-2000; 2000US-0679272.  
 XX 28-MAR-2001; 2001US-0822827.  
 XX (XUJ/) XU J.  
 XX (DILL/) DILLON D C.  
 XX (MITC/) MITCHAM J L.  
 XX (HARL/) HARLOCKER S L.  
 XX (JIAN/) JIANG Y.  
 XX (KALO/) KALOS M D.  
 XX (FANG/) FANGER G R.  
 XX (RETT/) RETTER M W.  
 XX (STOL/) STOLK J A.  
 XX (DAYC/) DAY C H.  
 XX (VEDV/) VEDVICK T S.  
 XX (CART/) CARTER D.  
 XX (LISK/) LI S X.  
 XX (WANG/) WANG A.  
 XX (SKEI/) SKEIKY Y A W.  
 XX (HEPL/) HEPLER W T.  
 XX (HEND/) HENDERSON R A.  
 XX (HURA/) HURAL J P D.  
 XX (MCNE/) MCNETTIL P D.  
 XX (HOUN/) HOUGHTON R L.  
 XX (DBAS/) Y DE BASOLS C V.  
 XX (FOYT/) FOY T M.  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;  
 PI McNeil PD, Houghton RL, Y De Basols CV, Foy TM;  
 DR WPI: 2003-352711/33.  
 XX New fusion protein comprising prostate-specific polypeptides, or its  
 PT immunogenic portions, useful for diagnosing, preventing and/or treating  
 PT cancer, particularly prostate cancer -  
 XX Example 15; SEQ ID NO 920; 85pp; English.  
 CC The invention describes a fusion protein comprising at least one amino  
 CC acid sequence of immunogenic portions of any of the 3 sequences not  
 CC defined in the specification, or sequences having at least 70 or 80 %  
 CC sequence identity to any one of the 3 sequences defined in the USPTO  
 CC web site, which is encoded by any of the 4 nucleotide sequences not  
 CC defined in the specification. The fusion protein, composition and  
 CC methods are useful for diagnosing, preventing and/or treating cancer,  
 CC particularly prostate cancer. The proteins are useful as markers to  
 CC indicate the presence or absence of cancer. This is the amino acid  
 CC sequence of a prostate cancer therapy associated protein.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from the US patent office at  
 CC seqdata.uspto.gov/sequence.html?DocID=US20020192763.  
 XX  
 SQ Sequence 318 AA;  
 Query Match 99.8%; Score 1614; DB 24; Length 318;  
 Best Local Similarity 99.7%; Pred. No. 5.9e-169;  
 Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDPNGNESATYFILIGLPGLEAQPWLAFPLCSLYLAVAGNTLIYVTEHSLHP 60

```

Db      2  MVDPNNGESSATYFLLIGLPGLEBAQFWLAPFLCSLYLAVLGNLTIIYVTRTHSLHEP 61
Qy      61  MYIFLCMLSGDILLISTSSMPKMLAIFWNSSTTIQFDACLLQIPIAHSLSGMESTVLLAM 120
Db      62  MYIFLCMLSGDILLISTSSMPKMLAIFWNSSTTIQFDACLLQIPIAHSLSGMESTVLLAM 121
Qy      121  AFDSYVAICHPLRATVTLTTRVTKIGVAAYVRGAALMAPLPVFIKQLPFCRSNIIISHSY 180
Db      122  AFDSYVAICHPLRATVTLTTRVTKIGVAAYVRGAALMAPLPVFIKQLPFCRSNIIISHSY 181
Qy      181  CLHODVMKLACDDIRVNVVYGLIYISAIGDSLISFSYLLIKTYLGLTRBAQAKAFG 240
Db      182  CLHODVMKLACDDIRVNVVYGLIYISAIGDSLISFSYLLIKTYLGLTRBAQAKAFG 241
Qy      241  TCVSHVCAVFIFYVPFGLSMVHRFSKRDSPLPVIIANTYLLVPPVLPNIYVGVTKKEI 300
Db      242  TCVSHVCAVFIFYVPFGLSMVHRFSKRDSPLPVIIANTYLLVPPVLPNIYVGVTKKEI 301
Qy      301  RQRIILRFHVATHASEP 317
Db      302  RQRIILRFHVATHASEP 318

```

```

RESULT 15
ABP81974
ID  ABP81974 standard; Protein, 318 AA.
XX
AC  ABP81974;
XX
DT  04-MAR-2003 (first entry)
XX
DE  Human G protein-coupled receptor l653440 protein SEQ ID NO:434.
XX

```

```

KM  G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KM  G protein-coupled receptor modulator; antibody; immune-related disease;
KM  growth-related disease; cell regeneration-related disease; AIDS; cancer;
KM  immunological-related cell proliferative disease; autoimmune disease;
KM  Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KM  osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KM  graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KM  psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KM  mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KM  hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
ulcer.
XX
OS  Homo sapiens.
XX
PN  W0200261087-A2.
XX
PD  08-AUG-2002.
XX
PF  19-DEC-2001; 2001WO-US50107.
XX
PR  19-DEC-2000; 2000US-257144P.
XX
RA  (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI  Burner GC, Roush CL, Brown JP;
XX
DR  WPI: 2003-045718/04.
XX
N-PSDB: ABZ42822.
XX

```

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases

Disclosure; Fig 1; 523pp; English.

The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular

```

CC  G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC  and (2) an isolated antibody having high specificity and high affinity
CC  or avidity for a particular GPCR. (1) can be used as GPCR modulators and
CC  in gene therapy. The antigenic peptides for GPCRs are useful in detecting
CC  an antibody against a particular GPCR, and in the production of specific
CC  antibodies. The peptides and antibodies are also useful for detecting the
CC  presence or absence of corresponding GPCRs. The antigenic peptides for
CC  GPCRs and antibodies are useful for diagnosing and designing drugs for
CC  treating immune-related diseases, growth-related diseases, cell
CC  regeneration-related disease, immunological-related cell proliferative
CC  diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC  atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC  osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC  inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC  disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC  anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC  loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC  hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC  any other disorder in which GPCRs are involved. The antibodies may be
CC  used in immunoassays and immunodiagnosis. ABZ4523 to ABZ42869 encode
CC  GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC  exemplification of the present invention.
XX
SQ  Sequence 318 AA;

```

```

Query Match      99.8%; Score 1614; DB 24; Length 318;
Best Local Similarity 99.7%; Pred. No. 5.9e-163;
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MVDPNNGESSATYFLLIGLPGLEBAQFWLAPFLCSLYLAVLGNLTIIYVTRTHSLHEP 60
Db      2  MVDPNNGESSATYFLLIGLPGLEBAQFWLAPFLCSLYLAVLGNLTIIYVTRTHSLHEP 61
Qy      61  MYIFLCMLSGDILLISTSSMPKMLAIFWNSSTTIQFDACLLQIPIAHSLSGMESTVLLAM 120
Db      62  MYIFLCMLSGDILLISTSSMPKMLAIFWNSSTTIQFDACLLQIPIAHSLSGMESTVLLAM 121
Qy      121  AFDSYVAICHPLRATVTLTTRVTKIGVAAYVRGAALMAPLPVFIKQLPFCRSNIIISHSY 180
Db      122  AFDSYVAICHPLRATVTLTTRVTKIGVAAYVRGAALMAPLPVFIKQLPFCRSNIIISHSY 181
Qy      181  CLHODVMKLACDDIRVNVVYGLIYISAIGDSLISFSYLLIKTYLGLTRBAQAKAFG 240
Db      182  CLHODVMKLACDDIRVNVVYGLIYISAIGDSLISFSYLLIKTYLGLTRBAQAKAFG 241
Qy      241  TCVSHVCAVFIFYVPFGLSMVHRFSKRDSPLPVIIANTYLLVPPVLPNIYVGVTKKEI 300
Db      242  TCVSHVCAVFIFYVPFGLSMVHRFSKRDSPLPVIIANTYLLVPPVLPNIYVGVTKKEI 301
Qy      301  RQRIILRFHVATHASEP 317
Db      302  RQRIILRFHVATHASEP 318

```

Search completed: February 9, 2004, 16:23:47  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 16:05:17 ; Search time 797 Seconds  
(without alignments)  
10621.628 Million cell updates/sec

Title: US-10-017-066a-1

Perfect score: 3136  
Sequence: 1 cagagagcgtctatttcagt.....aaaaaaaaaaaaaaaa 3136

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq.19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3136	100.0	3136	22	AAF83880 Human G-protein co
2	3136	100.0	3136	25	ABZ42822 Human G-protein co
3	2913.4	92.9	3104	24	ABK92202 Prostate cancer-as
4	2028.4	64.7	2061	22	AAH64175 Human prostate cdn
5	2028.4	64.7	2061	22	AAH93939 FLJ13581 cdna sequ
6	2028.4	64.7	2061	22	AAH18119 Human cdna sequenc
7	2028.4	64.7	2061	24	ABJ95546 Human FLJ13581 cdn
8	2028.4	64.7	2061	25	ACA59983 Prostate cancer th

Result 1	AAF83880	standard; cdna; 3136 BP.	ALIGNMENTS
ID	AAF83880	standard; cdna; 3136 BP.	
XX	XX	XX	
AC	AAF83880;		
DT	06-AUG-2001 (first entry)		
XX	XX	XX	
DE	Human G-protein coupled receptor, PHOR-1 encoding cdna.		
XX	XX	XX	
KW	G-protein-coupled receptor; prostate; cancer; PHOR-1; kidney; uterine; cervical; stomach; rectal; cytostatic; vaccine; cell function regulator;		
KW	human; prostate homologue of olfactory receptor-1; ss.		
XX	XX	XX	
OS	Homo sapiens.		
XX	XX	XX	
FT	Key	Location/Qualifiers	
FT	CDS	133..1086	
FT	FT	/tag= a	
FT	FT	/product= "PHOR-1"	
XX	XX	XX	
PN	WO200125434-A1.		
XX	XX	XX	
PD	12-APR-2001.		
XX	XX	XX	
PF	05-OCT-2000; 2000MO-US27543.		
XX	XX	XX	
PR	05-OCT-1999; 99US-0157902.		
XX	XX	XX	
PA	(UROC-) UROGENESYS INC.		
XX	XX	XX	

## ALIGNMENTS

Result 1	AAF83880	standard; cdna; 3136 BP.	ALIGNMENTS
ID	AAF83880	standard; cdna; 3136 BP.	
XX	XX	XX	
AC	AAF83880;		
DT	06-AUG-2001 (first entry)		
XX	XX	XX	
DE	Human G-protein coupled receptor, PHOR-1 encoding cdna.		
XX	XX	XX	
KW	G-protein-coupled receptor; prostate; cancer; PHOR-1; kidney; uterine; cervical; stomach; rectal; cytostatic; vaccine; cell function regulator;		
KW	human; prostate homologue of olfactory receptor-1; ss.		
XX	XX	XX	
OS	Homo sapiens.		
XX	XX	XX	
FT	Key	Location/Qualifiers	
FT	CDS	133..1086	
FT	FT	/tag= a	
FT	FT	/product= "PHOR-1"	
XX	XX	XX	
PN	WO200125434-A1.		
XX	XX	XX	
PD	12-APR-2001.		
XX	XX	XX	
PF	05-OCT-2000; 2000MO-US27543.		
XX	XX	XX	
PR	05-OCT-1999; 99US-0157902.		
XX	XX	XX	
PA	(UROC-) UROGENESYS INC.		
XX	XX	XX	

PI Raitano AB, Afar DEH, Jakobovits A, Paris M, Hubert RS;  
PI Mitchell SC, Saffran DC;  
DR WPI; 2001-367230/38.  
DR P-PSDB; AAB85002.

XX Novel gene designated PHOR-1, a G-protein-coupled receptor up-regulated  
PT in prostate cancer, useful as diagnostic marker and therapeutic target  
XX for cancers of prostate, kidney, uterus

PS Claim 1, Fig 1A-D, 139pp; English.

XX The invention relates to a novel G-protein-coupled receptor up-regulated  
CC in prostate cancer, termed PHOR-1. The encoding cDNA is contained in  
CC plasmid designated p101P3h1 deposited with ATCC as Accession No. P7A-312.  
CC PHOR-1 polypeptides and polynucleotides are useful for diagnosing the  
CC presence of cancer, especially prostate, kidney, uterine, cervical,  
CC stomach or rectal cancer by determining and comparing the level of the  
CC protein or mRNA expression in test and normal tissue samples.  
CC Pharmaceutical compositions comprising PHOR-1 is useful for treating  
CC cancer. PHOR-1 proteins are useful for identifying ligands and other  
CC agents and cellular constituents that binds to PHOR-1 gene product and  
CC for generating antibodies which are useful in diagnostic, prognostic and  
CC imaging methodologies and for the treatment of prostate cancer. Cell  
CC lines expressing PHOR-1 are useful for identifying protein-protein  
CC interactions mediated by PHOR-1. The present sequence represents a cDNA  
CC encoding the human PHOR-1 (prostate homologue of olfactory receptor-1).  
XX

Sequence 3136 BP; 864 A; 674 C; 622 G; 976 T; 0 other;

Query Match 100.0%; Score 3136; DB 22; Length 3136;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAGAGGCTGTAATTCAGTGCAGAGCTGCGACCTCTTCTGAGAGAAAGTGCAGAAAG 60  
DB 1 CAGAGAGGCTGTAATTCAGTGCAGAGCTGCGACCTCTTCTGAGAGAAAGTGCAGAAAG 60  
QY 61 GGGGTCACACATTCCTTCATACGATTGAGCTCTACCTGCGGTGCTGCACAGTTTC 120  
DB 61 GGGGTCACACATTCCTTCATACGATTGAGCTCTACCTGCGGTGCTGCACAGTTTC 120  
QY 121 AGCTTCCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
DB 121 AGCTTCCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
QY 181 ATAGGCTCTCTGCTGTTAGAGAGGCTCAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
DB 181 ATAGGCTCTCTGCTGTTAGAGAGGCTCAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
QY 241 TACCTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
DB 241 TACCTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
QY 301 CTGCGATGAGCCCATGAT 360  
DB 301 CTGCGATGAGCCCATGAT 360  
QY 361 ACCTCATGAGCCCAAAATGCTGCGCATCTTCTGCTCAATTCACATACCAATCAAGTTT 420  
DB 361 ACCTCATGAGCCCAAAATGCTGCGCATCTTCTGCTCAATTCACATACCAATCAAGTTT 420  
QY 421 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 421 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 481 CTGCTGAGCCATGAGCTTTGACCGCTATGAGCCATGCTGACCCACTGCGCATGAGCCACA 540  
DB 481 CTGCTGAGCCATGAGCTTTGACCGCTATGAGCCATGCTGACCCACTGCGCATGAGCCACA 540  
QY 541 GAACTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 541 GAACTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 601 CTGATGAGACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 601 CTGATGAGACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY 661 TCCCATTCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 661 TCCCATTCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
QY 721 AATGCT 780  
DB 721 AATGCT 780  
QY 781 TCCCTTCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 781 TCCCTTCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 AAGGATTTGGCAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
DB 841 AAGGATTTGGCAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 901 ATTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
DB 901 ATTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
QY 961 TTGGCAGATATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 961 TTGGCAGATATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 ACAAAGAGATTCAGACAGCGCATCTTGCATTCATGAGGACACAGCGCTTCAGAG 1080  
DB 1021 ACAAAGAGATTCAGACAGCGCATCTTGCATTCATGAGGACACAGCGCTTCAGAG 1080  
QY 1081 CCTAGGCT 1140  
DB 1081 CCTAGGCT 1140  
QY 1141 GTTAACTTTTGGAGAGACATTCAGAAATTTTCTTAAATTAATTAATTAATTAATTAATTAAT 1200  
DB 1141 GTTAACTTTTGGAGAGACATTCAGAAATTTTCTTAAATTAATTAATTAATTAATTAATTAAT 1200  
QY 1201 GATCCTTCAATATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
DB 1201 GATCCTTCAATATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
QY 1261 TGTTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
DB 1261 TGTTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1321 CTTTCATTTTACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
DB 1321 CTTTCATTTTACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
QY 1381 TGAGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
DB 1381 TGAGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
QY 1441 ATAAACAGAGAT 1500  
DB 1441 ATAAACAGAGAT 1500  
QY 1501 AACTCCCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
DB 1501 AACTCCCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
QY 1561 GAAATATATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
DB 1561 GAAATATATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
QY 1621 AAGAGTAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACT 1680  
DB 1621 AAGAGTAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACT 1680

```

QY 1681 TATGACCCCTGTTTCTTATTTATTTCTTATCAACCCCTTAAATTAGCCAAAGATTT 1740
Db 1681 TATGACCCCTGTTTCTTATTTATTTCTTATCAACCCCTTAAATTAGCCAAAGATTT 1740
QY 1741 ATTAGTACCCCTCATTTAGCCATGGGAAATGATGTCAGTGGGGATCAGGAATTA 1800
Db 1741 ATTAGTACCCCTCATTTAGCCATGGGAAATGATGTCAGTGGGGATCAGGAATTA 1800
QY 1801 TGGGGTCATACAGATATTAATAAAAAAAGACTTCATGCCAATCTCATATGA 1860
Db 1801 TGGGGTCATACAGATATTAATAAAAAAAGACTTCATGCCAATCTCATATGA 1860
QY 1861 TGGGAGAGACTGTAGAGACCAAGGATAGGGGTAGAGATTTCCAGAGCTTAC 1920
Db 1861 TGGGAGAGACTGTAGAGACCAAGGATAGGGGTAGAGATTTCCAGAGCTTAC 1920
QY 1921 ATTTCTAGAGAGATTAATTTCTTCTCATCATCCAGTGTGTATTTAGAAATTC 1980
Db 1921 ATTTCTAGAGAGATTAATTTCTTCTCATCATCCAGTGTGTATTTAGAAATTC 1980
QY 1981 CTGGCAACGAACCTCATGCTTTATCCCATAGCTATTTGCTTGTCTCAATT 2040
Db 1981 CTGGCAACGAACCTCATGCTTTATCCCATAGCTATTTGCTTGTCTCAATT 2040
QY 2041 GCCAATTACCTGTGTGGAAGAGTATTTCTAGGTTCAACCATATGAGAAATCTTA 2100
Db 2041 GCCAATTACCTGTGTGGAAGAGTATTTCTAGGTTCAACCATATGAGAAATCTTA 2100
QY 2101 TTTAGAAAGTCTGCATAGGGCTTATAGCAAGTATTTAATTTTAAAGTTCATAGTGA 2160
Db 2101 TTTAGAAAGTCTGCATAGGGCTTATAGCAAGTATTTAATTTTAAAGTTCATAGTGA 2160
QY 2161 TTTCTGTAGGCAAGTATAGGTTAGGAGCCACCATATATAGGAGATTTGAGAGGCGG 2220
Db 2161 TTTCTGTAGGCAAGTATAGGTTAGGAGCCACCATATATAGGAGATTTGAGAGGCGG 2220
QY 2221 TCTTGAAGATTAACATTTGGCTTTTGAAGTGTGACTGTAGCTGAAAGTGAAGCAATCTTC 2280
Db 2221 TCTTGAAGATTAACATTTGGCTTTTGAAGTGTGACTGTAGCTGAAAGTGAAGCAATCTTC 2280
QY 2281 AGGACCAATCTTTATTTGGGGCTTTGTGCAATATGAAACAGGACTTTGAGACCAAGAA 2340
Db 2281 AGGACCAATCTTTATTTGGGGCTTTGTGCAATATGAAACAGGACTTTGAGACCAAGAA 2340
QY 2341 GCAATCTGATAGGATGGAATCAGGATTTTCTTCTGAGGGGCTATTAACAAAGG 2400
Db 2341 GCAATCTGATAGGATGGAATCAGGATTTTCTTCTGAGGGGCTATTAACAAAGG 2400
QY 2401 TTAATAGGTTTCACTTCAACAGATATGACAAAGTGTAAACCAAGAACTCAATTC 2460
Db 2401 TTAATAGGTTTCACTTCAACAGATATGACAAAGTGTAAACCAAGAACTCAATTC 2460
QY 2461 AAATCTAAACATGATGATATATATGTAAGTATTTCTTCTTCAATCTCAGG 2520
Db 2461 AAATCTAAACATGATGATATATATGTAAGTATTTCTTCTTCAATCTCAGG 2520
QY 2521 TTCCCTGATATGATTTCTTAAACATGCTTCACTCCCTTTTGTATGATATCATATTT 2580
Db 2521 TTCCCTGATATGATTTCTTAAACATGCTTCACTCCCTTTTGTATGATATCATATTT 2580
QY 2581 GGAATGCTATTTAACTGATATTTGCTGTGATGATGAGGCAAGGCACTGTT 2640
Db 2581 GGAATGCTATTTAACTGATATTTGCTGTGATGATGAGGCAAGGCACTGTT 2640
QY 2641 TATTAATGATGATCTCTGTTCATCATGATGCTCTTGTGCTCATGATGATCCCC 2700
Db 2641 TATTAATGATGATCTCTGTTCATCATGATGCTCTTGTGCTCATGATGATCCCC 2700
QY 2701 AGCAAAAGTCCCTAGAAACATATATGCTTATGCTTGAACACGGTATTTTCAATCAACC 2760
Db 2701 AGCAAAAGTCCCTAGAAACATATATGCTTATGCTTGAACACGGTATTTTCAATCAACC 2760
QY 2761 TGATTCCTCTGTCTGAAACATATGACAGGCAATTTTCAAGCCTTCTTGTAGTGGGTA 2820

```

---

```

Db 2761 TGATTCCTCTGTCTGAAACATATGACAGGCAATTTTCAAGCCTTCTTGTAGTGGGTA 2820
QY 2821 TTAATTAATCTGAGCAATTAATCTTCAATGATGAGAGATGCAATTTCTATAC 2880
Db 2821 TTAATTAATCTGAGCAATTAATCTTCAATGATGAGAGATGCAATTTCTATAC 2880
QY 2881 CTGGCTCATTAACCCCTCCATGTGAGAGCTTTCATGTTGACATTAATGTGACTGGGA 2940
Db 2881 CTGGCTCATTAACCCCTCCATGTGAGAGCTTTCATGTTGACATTAATGTGACTGGGA 2940
QY 2941 AGCTATGTTTACACAGAGTAATACCAAGAACCTGATTTCTGAAAAAAGCTGCAGA 3000
Db 2941 AGCTATGTTTACACAGAGTAATACCAAGAACCTGATTTCTGAAAAAAGCTGCAGA 3000
QY 3001 GCCAAACCTCTGTCAATTTGCAATCCCACTTGTATTTGTATGAGAGAGTGTAAATGTA 3060
Db 3001 GCCAAACCTCTGTCAATTTGCAATCCCACTTGTATTTGTATGAGAGAGTGTAAATGTA 3060
QY 3061 AAATTAAGTACTATTTGTGTCAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3120
Db 3061 AAATTAAGTACTATTTGTGTCAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3120
QY 3121 AAAAAAAAAAAAAAAAAA 3136
Db 3121 AAAAAAAAAAAAAAAAAA 3136

RESULT 2
AB242822
ID AB242822 standard; DNA; 3136 BP.
XX
AC AB242822;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human G protein-coupled receptor Le53440 nucleotide SEQ ID NO:433.
XX
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor; modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related disease; cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.
XX
OS Homo sapiens.
XX
PN NC0200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001NC-0550107.
XX
PR 19-DEC-2000; 2000US-257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
PI Burner GC, Roush CL, Brown JP,
DR WPI; 2003-046718/04.
DR P-PDB; ABP81974.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating
PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
PT cancer or autoimmune diseases -
XX
PS Disclosure; Fig 1; 523pp; English.

```



Db 1681 TATGAGCCCTGTTTCTTATTAATTTCTATCAACCTTTATTTAGGCAAAATATTT 1740  
 Qy 1741 ATTATGACCTCATGTTAGCCATGGGAAATTTGATTCAGTGGGATCAGATTAATA 1800  
 Db 1741 ATTATGACCTCATGTTAGCCATGGGAAATTTGATTCAGTGGGATCAGATTAATA 1800  
 Qy 1801 TGGGGTCATACAGATTTAAAAATTTAAAAAAGACTTCATGCCCATCTCATATGA 1860  
 Db 1801 TGGGGTCATACAGATTTAAAAATTTAAAAAAGACTTCATGCCCATCTCATATGA 1860  
 Qy 1861 TGTGAGAAAGCTGTTAGAGACCAACAGGATAGTGGGTTAGAGATTTCCAGAGCTTAC 1920  
 Db 1861 TGTGAGAAAGCTGTTAGAGACCAACAGGATAGTGGGTTAGAGATTTCCAGAGCTTAC 1920  
 Qy 1921 ATTTTCTAGAGAGATTTAAATTTCTTCTCACTCATCCAGTGTGTATTTAGAAATTC 1980  
 Db 1921 ATTTTCTAGAGAGATTTAAATTTCTTCTCACTCATCCAGTGTGTATTTAGAAATTC 1980  
 Qy 1981 CTGGCAACAGAACTCATGCTTTATCCCACTAGCTATTTGCTTATGCTGCTCATTT 2040  
 Db 1981 CTGGCAACAGAACTCATGCTTTATCCCACTAGCTATTTGCTTATGCTGCTCATTT 2040  
 Qy 2041 GCCAATTACCTGTGTCTTGAAAGAGTATTTCTAGGTTCAACATTATGGAAGATTTCTTA 2100  
 Db 2041 GCCAATTACCTGTGTCTTGAAAGAGTATTTCTAGGTTCAACATTATGGAAGATTTCTTA 2100  
 Qy 2101 TTCAAGAAATCTGCATAGGCTTTATAGCAATTTATTTTAAAGTTCCATAGTGTGA 2160  
 Db 2101 TTCAAGAAATCTGCATAGGCTTTATAGCAATTTATTTTAAAGTTCCATAGTGTGA 2160  
 Qy 2161 TTCTGATAGGAGTGTAGTGAAGGACCACTGATATGATGGAGATGTGAAATGACAG 2220  
 Db 2161 TTCTGATAGGAGTGTAGTGAAGGACCACTGATATGATGGAGATGTGAAATGACAG 2220  
 Qy 2221 TCTTGAAAGATTAACATTTGAGTGTGATCTGTGAGTGTGAAAGTGTGAGGAACTTTC 2280  
 Db 2221 TCTTGAAAGATTAACATTTGAGTGTGATCTGTGAGTGTGAAAGTGTGAGGAACTTTC 2280  
 Qy 2281 AGGACCATGCTTTATTTGGGGCTTTGTGAGATGAAACAGGGACTTTGAGACCAAGAA 2340  
 Db 2281 AGGACCATGCTTTATTTGGGGCTTTGTGAGATGAAACAGGGACTTTGAGACCAAGAA 2340  
 Qy 2341 GCAATCTGACTTATGAGCATGGGATCAGGATTTTTCCTGAGGGGCTATTACCAAGG 2400  
 Db 2341 GCAATCTGACTTATGAGCATGGGATCAGGATTTTTCCTGAGGGGCTATTACCAAGG 2400  
 Qy 2401 TTATATGATTTTATCTTCAACAGATATGACAAAGTGTAAACCAAGAACTCAATATAC 2460  
 Db 2401 TTATATGATTTTATCTTCAACAGATATGACAAAGTGTAAACCAAGAACTCAATATAC 2460  
 Qy 2461 AAATATCAAAACATGATCATATATGTGTAGATTTTCTTTCTTTTCAATCCTCAGG 2520  
 Db 2461 AAATATCAAAACATGATCATATATGTGTAGATTTTCTTTCTTTTCAATCCTCAGG 2520  
 Qy 2521 TTCCCTGATATGATTTCTTAAACATGCTTATCCCTTTTGTATGATGATTCATTTT 2580  
 Db 2521 TTCCCTGATATGATTTCTTAAACATGCTTATCCCTTTTGTATGATGATTCATTTT 2580  
 Qy 2581 GGAATGCTTATTTAATCTTGAATTTGCTGTGACCTGATAGCCCATGAGGGGCTGTT 2640  
 Db 2581 GGAATGCTTATTTAATCTTGAATTTGCTGTGACCTGATAGCCCATGAGGGGCTGTT 2640  
 Qy 2641 TATATATGATGATCATCTCTGTTCATCATTTGATGATCTCTTTGCTCATATGATCCCC 2700  
 Db 2641 TATATATGATGATCATCTCTGTTCATCATTTGATGATCTCTTTGCTCATATGATCCCC 2700  
 Qy 2701 AGAAGATGCTTAAACATATATAGCTTATGCTTGAACACCGGATTTTTCATCAAAACC 2760  
 Db 2701 AGAAGATGCTTAAACATATATAGCTTATGCTTGAACACCGGATTTTTCATCAAAACC 2760  
 Qy 2761 TGAATTCCTTCTGTCTGAAACATATGACGAGCAATTTTCCAGCTTCTTGTAGTTGGGTA 2820

Db 2761 TGAATTCCTTCTGTCTGAAACATATGACGAGCAATTTTCCAGCTTCTTGTAGTTGGGTA 2820  
 Qy 2821 TTATTAATTTCTGGCATTACTTCCATGTGATGAGATGATGATGATTTCTATAC 2880  
 Db 2821 TTATTAATTTCTGGCATTACTTCCATGTGATGAGATGATGATGATTTCTATAC 2880  
 Qy 2881 CTGGCTCATTAACCCCTCCCATGTGACAGCTTTTCAATGTGACATTTAAATGTGACTGGGA 2940  
 Db 2881 CTGGCTCATTAACCCCTCCCATGTGACAGCTTTTCAATGTGACATTTAAATGTGACTGGGA 2940  
 Qy 2941 AGCTATGTGTACACAGATTAATCAACAGAACTGATTTTGTGAAAAAAGCTGACAG 3000  
 Db 2941 AGCTATGTGTACACAGATTAATCAACAGAACTGATTTTGTGAAAAAAGCTGACAG 3000  
 Qy 3001 GCCAAACCTCTGTCAATTTGCAATCTCCCATTTGATTTGTATGACAGGCAATTGATAGTA 3060  
 Db 3001 GCCAAACCTCTGTCAATTTGCAATCTCCCATTTGATTTGTATGACAGGCAATTGATAGTA 3060  
 Qy 3061 AAAATTAAGTACTATGTGTCAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3120  
 Db 3061 AAAATTAAGTACTATGTGTCAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3120  
 Qy 3121 AAAAAAAAAAAAAAAAAA 3136  
 Db 3121 AAAAAAAAAAAAAAAAAA 3136  
 RESULT 3  
 ABR92202  
 ID ABR92202 standard; DNA; 3104 BP.  
 AC ABR92202;  
 XX  
 DT 15-AUG-2002 (first entry)  
 XX  
 DE Prostate cancer-associated DNA sequence #88.  
 XX  
 KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;  
 KM gene therapy; gene; ds.  
 OS Mammalia.  
 XX  
 FN M0200230268-A2.  
 XX  
 PD 18-APR-2002.  
 XX  
 PE 12-OCT-2001; 2001MO-US32045.  
 XX  
 PR 13-OCT-2000; 2000US-0687576.  
 PR 08-DEC-2000; 2000US-0733288.  
 PR 08-DEC-2000; 2000US-0733742.  
 PR 24-JAN-2001; 2001US-263957P.  
 PR 16-MAR-2001; 2001US-276791P.  
 PR 16-MAR-2001; 2001US-276888P.  
 PR 06-APR-2001; 2001US-281922P.  
 PR 24-APR-2001; 2001US-286214P.  
 PR 30-APR-2001; 2001US-0847046.  
 PR 04-MAY-2001; 2001US-288589P.  
 XX  
 PA (BOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
 XX  
 DR WPI; 2002-471335/50.  
 DR P-PSDB; ABR61886.  
 XX  
 PT Detecting a prostate cancer-associated transcript in a cell in a  
 PT patient, useful for diagnosing prostate cancer (PC) or screening  
 PT modulators of PC, by determining if prostate cancer-associated genes  
 PT are expressed in a prostate tissue  
 XX  
 PS Claim 22; Page 372-373; 436pp; English.







```

Db      1800 TGGGGTCATACAGTATATAAAATT--AAAAAAAAGACTTCATGCCCAATCATATGA 1857
QY      1861 TGTGAGAGAACTGTATAGAGAGACCAACAGGGTAGGGGTAGAGATTTCCAGAGCTTAC 1920
Db      1858 TGTGAGAGAACTGTATAGAGAGACCAACAGGGTAGGGGTAGAGATTTCCAGAGCTTAC 1917
QY      1921 ATTTCTAGAGAGAGATTTATTTCTCTCATCTATCCAGAGTGTCTATTTAGAGATTTTC 1980
Db      1918 ATTTCTAGAGAGAGATTTATTTCTCTCATCTATCCAGAGTGTCTATTTAGAGATTTTC 1977
QY      1981 CTGGCAACAGAGCTCATGCTTTATATCCCACTAGTATGCTTATGCTGCTGCTCAATT 2040
Db      1978 CTGGCAACAGAGCTCATGCTTTATATCCCACTAGTATGCTTATGCTGCTGCTCAATT 2037
QY      2041 GCCAATTACTGCTGTCTTGGAGAGAGTATTTCTAGGTTCCACCATTTAGAGATTTCTTA 2100
Db      2038 GCCAATTACTGCTGTCTTGGAGAGAGTATTTCTAGGTTCCACCATTTAGAGATTTCTTA 2097
QY      2101 TTGAGAAAGTCTGATGAGGCTTATAGCAAGTATTTATTTTAAAGTTCCATAGGTGA 2160
Db      2098 TTGAGAAAGTCTGATGAGGCTTATAGCAAGTATTTATTTTAAAGTTCCATAGGTGT 2157
QY      2161 TTCTGATAGGCACTGAGTATGAGGAGCCACAGTTATGATGGAGATATGGAATGGCAGG 2220
Db      2158 TTCTGATAGGCACTGAGTATGAGGAGCCACAGTTATGATGGAGATATGGAATGGCAGG 2217
QY      2221 TCTTGAAGATACATAGGCTTTTGTAGTGTGACTGCTACTGGAAGTAGAGGAATCTTC 2280
Db      2218 TGTGAGAGATACATAGGCTTTTGTAGTGTGACTGCTACTGGAAGTAGAGGAATCTTC 2277
QY      2281 AGGACCACTGCTTATTTGGGGCTTTTGTGAGATATGAGACAGGACCTTTAGACAGAGAA 2340
Db      2278 AGGACCACTGCTTATTTGGGGCTTTTGTGAGATATGAGACAGGACCTTTAGACAGGAGAA 2337
QY      2341 GCATCTGATTTGGGCAATGAGGATTCAGGCAATTTTCTTCTGAGGGGCTATTTACCAAGGG 2400
Db      2338 GCATCTGATTTGGGCAATGAGGATTCAGGCAATTTTCTTCTGAGGGGCTATTTACCAAGGG 2397
QY      2401 TTATAGAGTTTCACTCTTCAACAGAGATATGACAAACAGTGTAAACCAAGAACTCAATTAC 2460
Db      2398 TTATAGAGTTTCACTCTTCAACAGAGATATGACAAACAGTGTAAACCAAGAACTCAATTAC 2457
QY      2461 AAATACTAAACATGTGATCATATATGTGTGAAGTTTCATTTCTTTTCAATCCTCAGG 2520
Db      2458 AAATACTAAACATGTGATCATATATGTGTGAAGTTTCATTTCTTTTCAATCCTCAGG 2517
QY      2521 TTCCCTGATATGAGATTCCTATAACAGTCTTCAATCCCTTTGTATGATGATATCATATTT 2580
Db      2518 TTCCCTGATATGAGATTCCTATAACAGTCTTCAATCCCTTTGTATGATGATATCATATTT 2577
QY      2581 GGAATGCTTATTTATCTTGTATTTGTGCTGAGCTGTAAAGCCCATGAGGCACTGTT 2640
Db      2578 GGAATGCTTATTTATCTTGTATTTGTGCTGAGCTGTAAAGCCCATGAGGCACTGTT 2637
QY      2641 TATATTTGATGATGATCTCTGTTCATCATTTGACTGTCTTTGCTCATCTTGAATCCCCC 2700
Db      2638 TATATTTGATGATGATCTCTGTTCATCATTTGACTGTCTTTGCTCATCTTGAATCCCCC 2697
QY      2701 AGCAAGTGTCTGAGACATATATGCTTATGCTTACACCGGTATTTTTCATCAAAAC 2760
Db      2698 AGCAAGTGTCTGAGACATATATGCTTATGCTTACACCGGTATTTTTCATCAAAAC 2757
QY      2761 TGATTTCTTCTGTCTGGAACACATAGCAGGCAATTTCCAGCCTTTCTTGTAGTTGGGTA 2820
Db      2758 TGATTTCTTCTGTCTGGAACACATAGCAGGCAATTTCCAGCCTTTCTTGTAGTTGGGTA 2817
QY      2821 TTATTAATTTCTGGCACTACTTCCATGATGAGTGAAGTGAACATGTCATATTTCTATAC 2880
Db      2818 TTATTAATTTTTAGCAATTTACTTCCATGATGAGTGAAGTGAACATGTCATATTTTATAC 2877
QY      2881 CTGGCTCATTAACCCCTCCCATGTGCAAGCCTTTCAATGTTGACATTAATGTGACTTGGGA 2940
Db      2878 CTGGCTCATTAACCCCTCCCATGTGCAAGCCTTTCAATGTTGACATTAATGTGACTTGGGA 2937

```

```

QY      2941 AGCTATGTTTACACAGAGTAAATCACC 2968
Db      2938 AGCTATGTTTACACAGAGTAAATTAAC 2965

RESULT 4
AAS64175
ID      AAS64175 standard; cDNA; 2061 BP.
XX
AC      AAS64175;
XX
DT      29-JAN-2002 (first entry)
XX
DE      Human prostate cDNA sequence #594.
XX
KW      Human; prostate cancer; ss; cytosolic; immunostimulant; tumour.
XX
OS      Homo sapiens.
XX
PN      WO200173032-A2.
XX
PD      04-OCT-2001.
XX
PF      27-MAR-2001; 2001WO-US09919.
XX
PR      27-MAR-2000; 2000US-0536857.
PR      09-MAY-2000; 2000US-0568100.
PR      12-MAY-2000; 2000US-0570737.
PR      13-JUN-2000; 2000US-0593793.
PR      27-JUN-2000; 2000US-0605783.
PR      10-AUG-2000; 2000US-0636215.
PR      29-AUG-2000; 2000US-0651236.
PR      06-SEP-2000; 2000US-0657279.
PR      02-OCT-2000; 2000US-0679426.
PR      10-OCT-2000; 2000US-0685166.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI      Ranger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI      Li SX, Wang A, Skeiky YW, Hepler WT, Henderson RA;
XX
DR      WPI; 2001-639232/73.
XX
PT      New human prostate-specific polypeptides and polynucleotides useful for
PT      the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS      Claim 1; Page 566-567; 579pp; English.
XX
CC      The invention relates to isolated prostate-specific
CC      polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC      antibodies raised against the polypeptides (or antigenic epitopes
CC      derived from them) and antigen-presenting cells expressing the
CC      polypeptides. The antibodies are useful for detecting the presence of
CC      cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC      the antigen-presenting cells are useful for stimulating and/or expanding
CC      T cells specific for a tumour protein, and for inhibiting the development
CC      of cancer especially prostate cancer. Compositions comprising the
CC      polynucleotide and/or polypeptide are useful for stimulating an immune
CC      response, and for treating cancer. The oligonucleotide is useful for
CC      detecting cancer. The present sequence is a prostate specific
CC      polynucleotide of the invention.
XX
SQ      Sequence 2061 BP; 612 A; 383 C; 399 G; 667 T; 0 other;

Query Match      64.7%; Score 2028.4; DB 22; Length 2061;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY      1029 GATTGACAGCGGATCTTGACATTTTCATGAGGACACAGCGCTTCAGAGCCCTAGGT 1088
Db      3 GATTGACAGCGGATCTTGACATTTTCATGAGGACACAGCGCTTCAGAGCCCTAGGT 62

```



DE FLJ13581 cDNA sequence.  
XX  
XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
XX cytostatic; gene therapy; metastasis; ss.  
XX  
XX Homo sapiens.  
XX  
XX MO200151633-A2.  
XX  
XX 19-JUL-2001.  
XX  
XX  
XX 16-JAN-2001; 2001WO-US01574.  
XX  
XX 14-JAN-2000; 2000US-0483672.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG,  
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skelky VM,  
PI Wang A, Mesgher MJ;  
XX  
XX WPI; 2001-425873/45.  
XX  
XX New polynucleotide encoding a prostate-specific protein, for  
PT diagnosing, monitoring and treating prostate cancer in a patient and  
PT for use in vaccines -  
XX  
XX  
XX Claim 1; Page 532-533; 543pp; English.  
XX  
XX The present invention describes polynucleotide sequences (I) which encode  
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
CC and can be used in vaccine production and gene therapy. (I), (II),  
CC antibodies to (II), fusion proteins comprising (II), and isolated  
CC T cells prepared using (I) or (II) are used to treat cancer in a patient.  
CC (I) and the antibodies are also used in the detection of cancer in a  
CC patient. The cancer that is diagnosed or treated is particularly  
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
CC (I) can be used for monitoring the progression of cancer in a patient.  
CC (I) and (II) can also be used to improve diagnostic and therapeutic  
CC methods for prostate cancer. They can indicate the level of metastasis  
CC as well as the prostate volume. AAH93357 to AAH93944 and AA01115 to  
CC AA01318 represent polynucleotide and amino acid sequences used in the  
CC exemplification of the present invention.  
XX  
XX  
XX Sequence 2061 BP; 612 A; 383 C; 399 G; 667 T; 0 other;  
SQ  
Query Match 64.7%; Score 2028.4; DB 22; Length 2061;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;  
QY 1029 GATTGACAGCGCATCTTCGACTTTTCATGTCGACACACAGCTTCAGAGCCCTAGGT 1088  
DB 3 GATTGACAGCGCATCTTCGACTTTTCATGTCGACACACAGCTTCAGAGCCCTAGGT 62  
QY 1089 GTCAGAGATCAACTCTTTCCATTCAGAGTCCTGATTCAGATTGTAATGTAAT 1148  
DB 63 GTCAGAGATCAACTCTTTCCATTCAGAGTCCTGATTCAGATTGTAATGTAAT 122  
QY 1149 TTTGAGAGACATATTCAGAAAAAAATTTCTTAATAAAAAATCACTAGATCCTTC 1208  
DB 123 TTTGAGAGACATATTCAGAAAAAAATTTCTTAATAAAAAATCACTAGATCCTTC 181  
QY 1209 AAATATGAAATGCTGTGGGAAATTCATTTTCAATATATATTTCTTCTTGTTCCT 1268  
DB 182 AAATATGAAATGCTGTGGGAAATTCATTTTCAATATATATTTCTTCTTGTTCCT 241  
QY 1269 TGTCTACAT 1388  
DB 242 TGTCTACAT 301  
QY 1329 TTTACATGACATGCAATATGAAATGCTTCTACTGATGCTTTACAGCATTCGAGATA 1388  
DB 302 TTTACATGACATGCAATATGAAATGCTTCTACTGATGCTTTACAGCATTCGAGATA 361

QY 1389 GAATGATACATCTAGAGAACATTTGCCAAAGGCTTAGACAGCGCAAGAAATTAACAC 1448  
DB 362 GAATGATACATCTAGAGAACATTTGCCAAAGGCTTAGACAGCGCAAGAAATTAACAC 421  
QY 1449 AGAAT 1508  
DB 422 AGAAT 481  
QY 1509 ACCAATTTGATCTCAGAAAAATGCTGTCTTCAAAATGACTTCAAGAGAAATTAAT 1568  
DB 482 ACCAATTTGATCTCAGAAAAATGCTGTCTTCAAAATGACTTCAAGAGAAATTAAT 541  
QY 1569 TTTTCTCTGAGACATAGCACTTAAGGGAAGATGGAAGTAAGCCTTGAAGAGTAC 1628  
DB 542 TTTTCTCTGAGACATAGCACTTAAGGGAAGATGGAAGTAAGCCTTGAAGAGTAC 601  
QY 1629 ATTACCTAGCTTAATGAAGTTGACACACTGTTCGAGAGTTTCAAGCATATAGGACC 1688  
DB 602 ATTACCTAGCTTAATGAAGTTGACACACTGTTCGAGAGTTTCAAGCATATAGGACC 661  
QY 1689 CTGTTTCTCTATTTAATTTCTTATCAACCTTAAATAGGCAAGATAATTAAGTAC 1748  
DB 662 CTGTTTCTCTATTTAATTTCTTATCAACCTTAAATAGGCAAGATAATTAAGTAC 721  
QY 1749 CCTCATTTGACATAGGGAATTTGATGTTCACTGGGATCACTGATTAATTAAGGGTCA 1808  
DB 722 CCTCATTTGACATAGGGAATTTGATGTTCACTGGGATCACTGATTAATTAAGGGTCA 781  
QY 1809 TACAGAT 1868  
DB 782 TACAGAT 839  
QY 1869 AACTGTTAGAGACCAACAGGAGTGGTGAAGATTTCCAGAGTCTTAATTTCTA 1928  
DB 840 AACTGTTAGAGACCAACAGGAGTGGTGAAGATTTCCAGAGTCTTAATTTCTA 899  
QY 1929 GAGAGGATTTAATTTCTTCACTCACTAGCTGATTAAGAAATTCCTGGCAAC 1988  
DB 900 GAGAGGATTTAATTTCTTCACTCACTAGCTGATTAAGAAATTCCTGGCAAC 959  
QY 1989 AGAATCTAGCTTTAATCCACTAGCTAATGCTTAATGCTGCTCAATTCGAATTA 2048  
DB 960 AGAATCTAGCTTTAATCCACTAGCTAATGCTTAATGCTGCTCAATTCGAATTA 1019  
QY 2049 CTTGCTGCTGGAAGAGTATTTCTAGGTCACCTTAAGTAAGTCTTATTCAGAA 2108  
DB 1020 CTTGCTGCTGGAAGAGTATTTCTAGGTCACCTTAAGTAAGTCTTATTCAGAA 1079  
QY 2109 GTCGTATAGGCTTTATAGCAATTTAATTTAAAGTTCATAGGATTCGATA 2168  
DB 1080 GTCGTATAGGCTTTATAGCAATTTAATTTAAAGTTCATAGGATTCGATA 1139  
QY 2169 GGCAGTGAAGTTAGGAGCCACCAATTAATGATGGAAGATAGCAAGCTTGAAG 2228  
DB 1140 GGCAGTGAAGTTAGGAGCCACCAATTAATGATGGAAGATAGCAAGCTTGAAG 1199  
QY 2229 ATAACTAGGCTTTGAGTGAATGCTAGCTGGAAGGGAATCTTCAGACCAT 2288  
DB 1200 ATAACTAGGCTTTGAGTGAATGCTAGCTGGAAGGGAATCTTCAGACCAT 1259  
QY 2289 GCTTTAATTTGGGCTTTGTGACATAGAAACAGGAATTTAGACACAGAAACCAATCTG 2348  
DB 1260 GCTTTAATTTGGGCTTTGTGACATAGAAACAGGAATTTAGACACAGAAACCAATCTG 1319  
QY 2349 ACTTAGGATGGAATCAGGCAATTTTGTCTGAGGGGCTTATCAAGGGTATAGG 2408  
DB 1320 ACTTAGGATGGAATCAGGCAATTTTGTCTGAGGGGCTTATCAAGGGTATAGG 1379  
QY 2409 TTTCAATCTTCAACAGATATGACACAGTGTAAACCAAGAAATCAATTAACAAATACTA 2468  
DB 1380 TTTCAATCTTCAACAGATATGACACAGTGTAAACCAAGAAATCAATTAACAAATACTA 1439

QY 2469 AAACATGATCATATATGTGTAGTTCATTTCTTTTCAATCTCTAGTTCCTCTA 2528  
 DB 1440 AAACATGATCATATATGTGTAGTTCATTTCTTTTCAATCTCTAGTTCCTCTA 1499  
 QY 2529 TATGATTCCTAATAACATGCTTTCATCCCTTTTGAATGATATCATATTTGGAAATGC 2588  
 DB 1500 TATGATTCCTAATAACATGCTTTCATCCCTTTTGAATGATATCATATTTGGAAATGC 1559  
 QY 2589 CTATTTAATCTGTGATTTGTGCTGACCTGTGAGCCCAAGAGGCACTGTTATTTATG 2648  
 DB 1560 CTATTTAATCTGTGATTTGTGCTGACCTGTGAGCCCAAGAGGCACTGTTATTTATG 1619  
 QY 2649 AATGTCATCTGTGTCATCATATGATCTGCTTTGCTCATATGATCCCAAGAGG 2708  
 DB 1620 AATGTCATCTGTGTCATCATATGATCTGCTTTGCTCATATGATCCCAAGAGG 1679  
 QY 2709 GCTTGAACATTAATATGCTTATGCTTGAACACCGGTTATTTTCATCAAACTGATTCCT 2768  
 DB 1680 GCTTGAACATTAATATGCTTATGCTTGAACACCGGTTATTTTCATCAAACTGATTCCT 1739  
 QY 2769 TCTGTCTGAAACATATGACCCAGGCAATTTTCAAGCCTTTGAGTTGGTATTTATA 2828  
 DB 1740 TCTGTCTGAAACATATGACCCAGGCAATTTTCAAGCCTTTGAGTTGGTATTTATA 1799  
 QY 2829 TTTGCGCATTAATCTTCCATATGATGAGTGAAGTGAATGCAATTTCTATACCTGCTCA 2888  
 DB 1800 TTTGCGCATTAATCTTCCATATGATGAGTGAAGTGAATGCAATTTCTATACCTGCTCA 1859  
 QY 2889 TAAACCCCTCCATGTCAGGCTTCTCATGTGACATTAATGTGACTTGGGAAAGCTATGT 2948  
 DB 1860 TAAACCCCTCCATGTCAGGCTTCTCATGTGACATTAATGTGACTTGGGAAAGCTATGT 1919  
 QY 2949 GTTACACAGAGTAATATCAGCAAGAGCTGATTTCTGAAAAAACTGTGACAGCCAAAC 3008  
 DB 1920 GTTACACAGAGTAATATCAGCAAGAGCTGATTTCTGAAAAAACTGTGACAGCCAAAC 1979  
 QY 3009 TCTGTCTTGTGCACTCCCACTTGTATTTGTACAGGAGCTGTGATTAAGTAAATAATA 3068  
 DB 1980 TCTGTCTTGTGCACTCCCACTTGTATTTGTACAGGAGCTGTGATTAAGTAAATAATA 2039  
 QY 3069 GTACTATTTGTGTCAAG 3084  
 DB 2040 GTACTATTTGTGTCAAG 2055  
 RESULT 6  
 AAH18119  
 ID AAH18119 standard; cDNA; 2061 BP.  
 AC AAH18119;  
 DT 26-JUN-2001 (first entry)  
 DE Human cDNA sequence SEQ ID NO:17984.  
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 OS Homo sapiens.  
 XX EPI074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JF-0248036.  
 XX 27-AUG-1999; 99JF-0300253.  
 XX 11-JAN-2000; 2000JP-0118776.  
 XX 02-MAY-2000; 2000JP-0183767.  
 XX 03-JUN-2000; 2000JP-0241899.  
 PA (HELI-) HELIX RES INST.  
 XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-118749/34.  
 DR  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 PS  
 PS Claim 8; SEQ ID 17984; 2537bp + CD ROM; English.  
 XX

CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH1628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH35893 represent human amino acid sequences; and AAH1629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX

SQ Sequence 2061 BP; 612 A; 383 C; 399 G; 667 T; 0 other;

Query Match 64.7%; Score 2028.4; DB 22; Length 2061;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1029 GATTCGACAGGCGACCTTCCTGACCTTTTCATGTGGCCACACAGGCTTCAAGCCCTAAGT 1088  
 DB 3 GATTCGACAGGCGACCTTCCTGACCTTTTCATGTGGCCACACAGGCTTCAAGCCCTAAGT 62  
 QY 1089 GTCAGTATCAAACTTCTTTTCCATTCAGAGTCTCTGATTCAGATTTTAAGTTAATCAT 1148  
 DB 63 GTCAGTATCAAACTTCTTTTCCATTCAGAGTCTCTGATTCAGATTTTAAGTTAATCAT 122  
 QY 1149 TTGGAAGACAGTATTCAGAAAAAAATTTCTTAATAAAAATACACTCAGATCCTTC 1208  
 DB 123 TTGGAAGACAGTATTCAGAAAAAAATTTCTTAATAAAAATACACTCAGATCCTTC 181  
 QY 1209 AAATATGAACCTGTGGGGAATCTCCATTTTTCATATATATTTCTTCTTGTTCCT 1268  
 DB 182 AAATATGAACCTGTGGGGAATCTCCATTTTTCATATATATTTCTTCTTGTTCCT 241  
 QY 1269 TGCTACATATATATTAATACCTGACTAGTGTGTGGAGGCTTATTAATCTTTCAT 1328  
 DB 242 TGCTACATATATATTAATACCTGACTAGTGTGTGGAGGCTTATTAATCTTTCAT 301  
 QY 1329 TTTACCATGCAAGTCCAAATCTAACTGCTTCTACTGATGTTTACAGCAATTCGAGTAA 1388  
 DB 302 TTTACCATGCAAGTCCAAATCTAACTGCTTCTACTGATGTTTACAGCAATTCGAGTAA 361  
 QY 1389 GAATGTATCATCTAGAGACATTTGGCAAGGCTTACAGCGCAAGGAAATTAACAC 1448  
 DB 362 GAATGTATCATCTAGAGACATTTGGCAAGGCTTACAGCGCAAGGAAATTAACAC 421  
 QY 1449 AGAATATATTAATGAATGATATCTAGCTTAATAACTATTAACCTCTCTTCAAGACTCCCA 1508  
 DB 422 AGAATATATTAATGAATGATATCTAGCTTAATAACTATTAACCTCTCTTCAAGACTCCCA 481

QY 1509 ACCACATGGATCTCAGAAAAAGCTGCTTCAAAAATGACTTCTACAGAGAGAAATAT 1568  
 DB 482 ACCACATGGATCTCAGAAAAAGCTGCTTCAAAAATGACTTCTACAGAGAGAAATAT 541  
 QY 1569 TTTTCTCTGACACTGACCTTAAAGGGAGATTGAAAGTAAAGCTTTGAAAAGATAG 1628  
 DB 542 TTTTCTCTGACACTGACCTTAAAGGGAGATTGAAAGTAAAGCTTTGAAAAGATAG 601  
 QY 1629 ATTACCTCAGTTAATGAAAGTGAACACATGCTGTCGAGAGTTTCAAGACATATGAGCC 1688  
 DB 602 ATTACCTCAGTTAATGAAAGTGAACACATGCTGTCGAGAGTTTCAAGACATATGAGCC 661  
 QY 1689 CTGTTTTCCTAATTAATTTCTTAACAACCTTTAATTAGGCAAGATATTAATAGTAC 1748  
 DB 662 CTGTTTTCCTAATTAATTTCTTAACAACCTTTAATTAGGCAAGATATTAATAGTAC 721  
 QY 1749 CCTCATTTAGACATGAGGAAATTTGATGTCAGTGGGATCAGTGAATTAATGAGGCTCA 1808  
 DB 722 CCTCATTTAGACATGAGGAAATTTGATGTCAGTGGGATCAGTGAATTAATGAGGCTCA 781  
 QY 1809 TACAAGTATAAAATTTAAAAAAAAGACTTCATGCCCAATCTCATATGATGAGAG 1868  
 DB 782 TACAAGTATAAAATTTAAAAAAAAGACTTCATGCCCAATCTCATATGATGAGAG 839  
 QY 1869 AACTGTTAGAGACCAACAGGGTAGTGGTTAGAGATTTCAGAGTCTTACATTTCTA 1928  
 DB 840 AACTGTTAGAGACCAACAGGGTAGTGGTTAGAGATTTCAGAGTCTTACATTTCTA 899  
 QY 1929 GAGGAGGTATTAATTTCTCTCATCANCCAGTGTGATTTTAAAGATTTTCCCTGGCAAC 1988  
 DB 900 GAGGAGGTATTAATTTCTCTCTCATCANCCAGTGTGATTTTAAAGATTTTCCCTGGCAAC 959  
 QY 1989 AGAAGCTCATGAGCTTTAATCCCACTAGCTATTGCTTATTTGCTGCTCCAAATGCAATTA 2048  
 DB 960 AGAAGCTCATGAGCTTTAATCCCACTAGCTATTGCTTATTTGCTGCTCCAAATGCAATTA 1019  
 QY 2049 CCTGTGTCTTGAAGAAGGATTTCTAGGTCACCATTAATGAGAGATTTCTTATTCAGAA 2108  
 DB 1020 CCTGTGTCTTGAAGAAGGATTTCTAGGTCACCATTAATGAGAGATTTCTTATTCAGAA 1079  
 QY 2109 GTCTGCATAGGGCTTATAGCAAGTATTAATTTTAAAAAGTTTCCATAGGATCTGATA 2168  
 DB 1080 GTCTGCATAGGGCTTATAGCAAGTATTAATTTTAAAAAGTTTCCATAGGATCTGATA 1139  
 QY 2169 GGCAGTGAAGTTAGGAGCCACAGTATGATGGAGATGAGATGGCAGGCTTGAAG 2228  
 DB 1140 GGCAGTGAAGTTAGGAGCCACAGTATGATGGAGATGAGATGGCAGGCTTGAAG 1199  
 QY 2229 ATTAACATTTGGCTTTGAGTGTGACTGCTAGCTGGAAGTGAAGGAACTTCAAGACAT 2288  
 DB 1200 ATTAACATTTGGCTTTGAGTGTGACTGCTAGCTGGAAGTGAAGGAACTTCAAGACAT 1259  
 QY 2289 GCTTTATTTGGGCTTTGAGTGTGACTGCTAGCTGGAAGTGAAGGAACTTCAAGACAT 2348  
 DB 1260 GCTTTATTTGGGCTTTGAGTGTGACTGCTAGCTGGAAGTGAAGGAACTTCAAGACAT 1319  
 QY 2349 ACTTAGGCATGGGATTCAGGCAATTTTGTCTTCTGAGGGGCTATTACAGGGCTTAATAGG 2408  
 DB 1320 ACTTAGGCATGGGATTCAGGCAATTTTGTCTTCTGAGGGGCTATTACAGGGCTTAATAGG 1379  
 QY 2409 TTTCATCTTCAACAGGATATGACACAGTGTAAACAGAAATCAATTAACAATCTA 2468  
 DB 1380 TTTCATCTTCAACAGGATATGACACAGTGTAAACAGAAATCAATTAACAATCTA 1439  
 QY 2469 AAACATGTGATCATATATGTGGTAAGTTTCAATTTCTTTCATCTCAGGTTCCCTGA 2528  
 DB 1440 AAACATGTGATCATATATGTGGTAAGTTTCAATTTCTTTCATCTCAGGTTCCCTGA 1499  
 QY 2529 TATGATTCCTATATACATGCTTTCATGCCCTTTGTATGATGATTCATATTTGGAATGC 2588  
 DB 1500 TATGATTCCTATATACATGCTTTCATGCCCTTTGTATGATGATTCATATTTGGAATGC 1559  
 QY 2589 CTATTTAATACTTGTATTTGTCTGTGACGTGAAGCCCATGAGGGCACTGTTTATATG 2648

DB 1560 CTATTTAATACTTGTATTTTCTGTCTGACCTTAAGCCCATGAGGGCACTGTATTAATG 1619  
 QY 2649 AATGTCATCTCTGTTCATCATATTGACTGCTTGTCTCATCATTTGAATCCCCAGCAAGT 2708  
 DB 1620 AATGTCATCTCTGTTCATCATATTGACTGCTTGTCTCATCATTTGAATCCCCAGCAAGT 1679  
 QY 2709 GCCTAGAACATATATAGCTTATAGCTTGAACACCGGTATTTTTCATCAAACTGATTCCT 2768  
 DB 1680 GCCTAGAACATATATAGCTTATAGCTTGAACACCGGTATTTTTCATCAAACTGATTCCT 1739  
 QY 2769 TCTGTCTGACACATAGCAGGCAATTTCCAGCTTCTTGTAGTGGTATTAATAA 2828  
 DB 1740 TCTGTCTGACACATAGCAGGCAATTTCCAGCTTCTTGTAGTGGTATTAATAA 1799  
 QY 2829 TTTGCGCATTAATCTTCATATGATGAGTGAAGTGAACATGTCGAATTTCTATACCTGCTCA 2888  
 DB 1800 TTTGCGCATTAATCTTCATATGATGAGTGAAGTGAACATGTCGAATTTCTATACCTGCTCA 1859  
 QY 2889 TAAACCCCTCCATAGTGAAGCTTTGATGATGATTAATGATGATGAGGAAAGCTATGT 2948  
 DB 1860 TAAACCCCTCCATAGTGAAGCTTTGATGATGATTAATGATGATGAGGAAAGCTATGT 1919  
 QY 2949 GTTACACAGATTAATCACACAGAGCCTGATTTCTGAATAAACTGTGACAGCCAAAC 3008  
 DB 1920 GTTACACAGATTAATCACACAGAGCCTGATTTCTGAATAAACTGTGACAGCCAAAC 1979  
 QY 3009 TCTGTATTTGCAATCCCATCTGTATTTGTAGAGAGCAGTGGATTAAGTAAAAATTA 3068  
 DB 1980 TCTGTATTTGCAATCCCATCTGTATTTGTAGAGAGCAGTGGATTAAGTAAAAATTA 2039  
 QY 3069 GTACTATTGTGTCAAG 3084  
 DB 2040 GTACTATTGTGTCAAG 2055

RESULT 7  
 ABL95546  
 ID ABL95546 standard; cDNA; 2061 BP.  
 XX  
 AC ABL95546;  
 XX  
 DT 19-JUL-2002 (first entry)  
 XX  
 DE Human FLJ13581 cDNA sequence SEQ ID NO 917.  
 XX  
 KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;  
 gene therapy; gene; ss.  
 OS Homo sapiens.  
 XX  
 PN US2002022248-A1.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 12-JAN-2001; 2001US-0759143.  
 XX  
 XX 25-FEB-1997; 97US-080609.  
 PR 01-AUG-1997; 97US-0904804.  
 PR 09-FEB-1998; 98US-0020956.  
 PR 25-FEB-1998; 98US-0030607.  
 PR 14-JUL-1998; 98US-0115453.  
 PR 22-SEP-1998; 98US-0159612.  
 PR 15-JAN-1999; 99US-0232149.  
 PR 09-APR-1999; 99US-0288946.  
 PR 13-JUL-1999; 99US-0352616.  
 PR 12-NOV-1999; 99US-0439313.  
 PR 18-NOV-1999; 99US-0443686.  
 PR 14-JAN-2000; 2000US-0483672.  
 PR 27-MAR-2000; 2000US-0536857.  
 PR 09-MAY-2000; 2000US-0568100.  
 PR 12-MAY-2000; 2000US-0570737.  
 PR 13-JUN-2000; 2000US-0593793.

[illegible]

```

Db 1440 AACATGATCATATATATGTAAGTTCATTTCTTTTCATTCCTGAGTCCCTGA 1499
Qy 2529 TATGATTCCTTAATACATGCTTTCATCCCTTTTGAATGATATCATTTGGAAATGC 2588
Db 1500 TATGATTCCTTAATACATGCTTTCATCCCTTTTGAATGATATCATTTGGAAATGC 1559
Qy 2589 CTATTTAATCTTGATTTTGTCTGCTGACAGTGAAGCCATGAGGACAGTTTATATG 2648
Db 1560 CTATTTAATCTTGATTTTGTCTGCTGACAGTGAAGCCATGAGGACAGTTTATATG 1619
Qy 2649 AATGATCTCTGTTTATCATGATGCTGCTTTTCTCATCATTAATCCCCAGCAAGT 2708
Db 1620 AATGATCTCTGTTTATCATGATGCTGCTTTTCTCATCATTAATCCCCAGCAAGT 1679
Qy 2709 GCCTAGACATATATAGTGTATGCTTGAACACCGGTTATTTTCAATCAAACTGATTCCT 2768
Db 1680 GCCTAGACATATATAGTGTATGCTTGAACACCGGTTATTTTCAATCAAACTGATTCCT 1739
Qy 2769 TCTGTCTGAAACATATAGCCAGCAATTTTCCAGCCTTTGAGTTGGGATTAATATA 2828
Db 1740 TCTGTCTGAAACATATAGCCAGCAATTTTCCAGCCTTTGAGTTGGGATTAATATA 1799
Qy 2829 TCTGTGCGCATTTACTTCCAAATGAGTGAAGAGTGAAGTGCATTTCTTACTGCTTCA 2888
Db 1800 TCTGTGCGCATTTACTTCCAAATGAGTGAAGAGTGAAGTGCATTTCTTACTGCTTCA 1859
Qy 2889 TAAACCCCTCCCATGTCAGCCCTTTTCTGATGTCATTAATGTCAGTGGAAAGCTATGT 2948
Db 1860 TAAACCCCTCCCATGTCAGCCCTTTTCTGATGTCATTAATGTCAGTGGAAAGCTATGT 1919
Qy 2949 GTTACACAGATTAATACCAAGAGCTGATTTCTGAAAAAATCTGTGACAGCCAAAC 3008
Db 1920 GTTACACAGATTAATACCAAGAGCTGATTTCTGAAAAAATCTGTGACAGCCAAAC 1979
Qy 3009 TCTGTCAATTTGCAACTCCCACTTGTATTTGTAGAGGACGTTGATTAAGTGAATAATA 3068
Db 1980 TCTGTCAATTTGCAACTCCCACTTGTATTTGTAGAGGACGTTGATTAAGTGAATAATA 2039
Qy 3069 GTACTATTGTGTCAAG 3084
Db 2040 GTACTATTGTGTCAAG 2055

```

## RESULT 8

ACAS9983

ID ACAS9983 standard; cDNA; 2061 BP.

XX ACAS9983;

XX 10-JUN-2003 (first entry)

XX Prostate cancer therapy associated cDNA #664.

XX Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;

XX Immunogen; cancer; prostate specific antigen; PSA;

XX Prostatic acid phosphatase; PAP; prostate specific membrane antigen;

XX BSM; gene; ss.

XX Homo sapiens.

XX US2002192763-A1.

XX 19-DEC-2002.

XX 29-JUN-2001; 2001US-0895793.

XX 17-APR-2000; 2000US-157455P.

XX 04-OCT-2000; 2000US-0679272.

XX 28-MAR-2001; 2001US-08282827.

XX (XUJ/) XU J.

XX (DILL/) DILLON D C.

XX (MITC/) MITCHAM J L.

```

PA (HARL/) HARLOCKER S L.
PA (JIAN/) JIANG Y.
PA (KALO/) KALOS M D.
PA (FANG/) FANGER G R.
PA (RETT/) RETTER M W.
PA (STOL/) STOLK J A.
PA (DAYC/) DAY C H.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (LISX/) LI S X.
PA (WANG/) WANG A.
PA (SKEI/) SKEIKY Y A W.
PA (HEPL/) HEPLER W T.
PA (HEND/) HENDERSON R A.
PA (HURA/) HURAL J.
PA (MCNE/) MCNEILL P D.
PA (HOUO/) HOUGHTON R L.
PA (DBAS/) Y DE BASOLS C V.
PA (FOYT/) FOY T M.

```

```

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;
PI Mcneill PD, Houghton RL, Y De Basols CV, Foy TM;
DR WPI; 2003-352711/33.

```

```

PT New fusion protein comprising prostate-specific polypeptides, or its
PT immunogenic portions, useful for diagnosing, preventing and/or treating
PT cancer, particularly prostate cancer.
PS Example 15; SEQ ID NO 917; 85bp; English.

```

The invention describes a fusion protein comprising at least one amino acid sequence of immunogenic portions of any of the 3 sequences not defined in the specification, or sequences having at least 70 or 90 % sequence identity to any one of the 3 sequences defined in the USPTO web site, which is encoded by any of the 4 nucleotide sequences not defined in the specification. The fusion protein, composition and methods are useful for diagnosing, preventing and/or treating cancer, particularly prostate cancer. The proteins are useful as markers to indicate the presence or absence of cancer. This sequence represents a prostate cancer therapy associated cDNA.

Note: The sequence data for this patent did not form part of the CC printed specification, but was obtained in electronic format directly from the US patent office at [seqdata.uspto.gov/sequence.html?docid=US20020192763](http://seqdata.uspto.gov/sequence.html?docid=US20020192763).

XX Sequence 2061 BP; 612 A; 383 C; 399 G; 667 T; 0 other;

```

Qy Query Match 64.7%; Score 2028.4; DB 25; Length 2061;
Qy Best Local Similarity 99.8%; Pred. No. 0;
Qy Matches 2052; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

```

```

Qy 1029 GATTGACAGGAGATCTTGTGACCTTTCATGTCAGAGTCTGATTCAGATTTAATGTAACAT 1148
Db 3 GATTGACAGGAGATCTTGTGACCTTTCATGTCAGAGTCTGATTCAGATTTAATGTAACAT 122
Qy 1089 GTCAGTATCAAACTTCTTTTCCATTCAGAGTCTGATTCAGATTTAATGTAACAT 1148
Db 63 GTCAGTATCAAACTTCTTTTCCATTCAGAGTCTGATTCAGATTTAATGTAACAT 122
Qy 1149 TTGGAAGACAGATTTTGAAGAAAAAATTTCCCTTAATATAAATACACACAGTCTTTC 1208
Db 123 TTGGAAGACAGATTTTGAAGAAAAAATTTCCCTTAAT-AAAAATACACACAGTCTTTC 161
Qy 1209 AAATATGAACCTGTTGGGAATCTCCATTTTTCATATATATATTTCTTTGTTTCT 1268
Db 182 AAATATGAACCTGTTGGGAATCTCCATTTTTCATATATATATTTCTTTGTTTCT 241
Qy 1269 TGCTACATATATATATATATATATATATATATATATATATATATATATATATATATAT 1328
Db 242 TGCTACATATATATATATATATATATATATATATATATATATATATATATATATATAT 301

```



QY 1329 TTACCATGAGTCCAAATCTAAAGCTGCTGACGATGTTTACAGATCTTGAGATA 1388  
 Db 302 TTACCATGAGTCCAAATCTAAAGCTGCTGACGATGTTTACAGATCTTGAGATA 361  
 QY 1389 GAATGATCATGTAGAGAACTTTGCCAAGGCTTACAGAGGAAAGAAATTAACAC 1448  
 Db 362 GAATGATCATGTAGAGAACTTTGCCAAGGCTTACAGAGGAAAGAAATTAACAC 421  
 QY 1449 AGAATATATTAAGATAGATATCTAGCTTAAATCTAACTTCTCTGAACTCCCA 1508  
 Db 422 AGAATATATTAAGATAGATATCTAGCTTAAATCTAACTTCTCTGAACTCCCA 481  
 QY 1509 ACCACATTTGATCTCAGAAAAATGCTGCTTCAAAATGACTTTACAGAGAGAAATAT 1568  
 Db 482 ACCACATTTGATCTCAGAAAAATGCTGCTTCAAAATGACTTTACAGAGAGAAATAT 541  
 QY 1569 TTTTCTCTGAGACATGACATCTTAAGAGGAAAGATTGGAAGTAAGCTTGAAGAGTAC 1628  
 Db 542 TTTTCTCTGAGACATGACATCTTAAGAGGAAAGATTGGAAGTAAGCTTGAAGAGTAC 601  
 QY 1629 ATTACCTAGCTTAATGAAGTGAACACATGTTCTGAGATTTTCAAGATATGAGAC 1688  
 Db 602 ATTACCTAGCTTAATGAAGTGAACACATGTTCTGAGATTTTCAAGATATGAGAC 661  
 QY 1689 CTGTTTCTCTAATTAATTTTCTTATCAACCTTTAATAGGCAAGATATTTAGTAC 1748  
 Db 662 CTGTTTCTCTAATTAATTTTCTTATCAACCTTTAATAGGCAAGATATTTAGTAC 721  
 QY 1749 CCGATTTGAGCAATGGGAAATTTGATGTTGAGGAGGATCAGTGAATTAAGGGGTCA 1808  
 Db 722 CCGATTTGAGCAATGGGAAATTTGATGTTGAGGAGGATCAGTGAATTAAGGGGTCA 781  
 QY 1809 TACAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1868  
 Db 782 TACAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 839  
 QY 1869 AACTGTTAGAGAGACCAACAGAGGATGAGTTAGAGATTTCCAGAGTCTTACATTTCTA 1928  
 Db 840 AACTGTTAGAGAGACCAACAGAGGATGAGTTAGAGATTTCCAGAGTCTTACATTTCTA 899  
 QY 1929 GAGAGGATTAATTTCTTCTCACTCATCCAGTGTGATTTAGAGATTTCCGGAAC 1988  
 Db 900 GAGAGGATTAATTTCTTCTCACTCATCCAGTGTGATTTAGAGATTTCCGGAAC 959  
 QY 1989 AGAATCTATGCTTTAATCCCACTAGCTATTTGCTGCTGCTGCTGCTGCTGCTGCT 2048  
 Db 960 AGAATCTATGCTTTAATCCCACTAGCTATTTGCTGCTGCTGCTGCTGCTGCTGCT 1019  
 QY 2049 CCTGATCTTGAAGAGATGTTCTAGTTCAACCTATTAAGAGATTTCTTATTCGAAA 2108  
 Db 1020 CCTGATCTTGAAGAGATGTTCTAGTTCAACCTATTAAGAGATTTCTTATTCGAAA 1079  
 QY 2109 GTCTGATAGGCTTATAGCAAGTATTTATTTTAAAGTCCATAGTATTCGATA 2168  
 Db 1080 GTCTGATAGGCTTATAGCAAGTATTTATTTTAAAGTCCATAGTATTCGATA 1139  
 QY 2169 GGCAGTATGTTAGAGGACCAACGTTATGATGGAAGTATGGAAGGAGCTTGAAG 2228  
 Db 1140 GGCAGTATGTTAGAGGACCAACGTTATGATGGAAGTATGGAAGGAGCTTGAAG 1199  
 QY 2229 ATAACATTTGCTTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2288  
 Db 1200 ATAACATTTGCTTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259  
 QY 2289 GCTTTATTTGGGCTTTTGTGAGATGGAACAGGAGCTTTGAGACCAAGAAACATCTG 2348  
 Db 1260 GCTTTATTTGGGCTTTTGTGAGATGGAACAGGAGCTTTGAGACCAAGAAACATCTG 1319  
 QY 2349 ACTTAGGATGGAATAGGATTTTGTGCTGAGAGGCTTATTAACAGGCTTATAG 2408  
 Db 1320 ACTTAGGATGGAATAGGATTTTGTGCTGAGAGGCTTATTAACAGGCTTATAG 1379

QY 2409 TTTGATCTTCAACAGATATGACACAGCTTTTACCAAGAACTCAATTAACATCTA 2468  
 Db 1380 TTTGATCTTCAACAGATATGACACAGCTTTTACCAAGAACTCAATTAACATCTA 1439  
 QY 2469 AAACATGATATATATGTTAGTTTCA TTTCTTTTCAATCTCAGGTCCTGTA 2528  
 Db 1440 AAACATGATATATATGTTAGTTTCA TTTCTTTTCAATCTCAGGTCCTGTA 1499  
 QY 2529 TATGATTTCTTATACATGCTTCAATCCCTTTTGTATGATGATATTTTGAATGC 2588  
 Db 1500 TATGATTTCTTATACATGCTTCAATCCCTTTTGTATGATGATATTTTGAATGC 1559  
 QY 2589 CTATTTATACCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2648  
 Db 1560 CTATTTATACCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619  
 QY 2649 AATGATCTCTGTTCAATGATGATGCTCTTGTGCTATGATGATGATGATGATGAT 2708  
 Db 1620 AATGATCTCTGTTCAATGATGATGCTCTTGTGCTATGATGATGATGATGATGAT 1679  
 QY 2709 GCTTGAATATATAGTGTATAGCTTATAGTGAACCGGTTATTTTCAATCAACCTGAT 2768  
 Db 1680 GCTTGAATATATAGTGTATAGCTTATAGTGAACCGGTTATTTTCAATCAACCTGAT 1739  
 QY 2769 TCTGCTGGAACATATAGTGTATAGCTTATAGTGAACCGGTTATTTTCAATCAACCT 2828  
 Db 1740 TCTGCTGGAACATATAGTGTATAGCTTATAGTGAACCGGTTATTTTCAATCAACCT 1799  
 QY 2829 TTTGAGCCTTATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2888  
 Db 1800 TTTGAGCCTTATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1859  
 QY 2889 TAAACCTTCCCATGAGGCTTATGATGATGATGATGATGATGATGATGATGATGAT 2948  
 Db 1860 TAAACCTTCCCATGAGGCTTATGATGATGATGATGATGATGATGATGATGATGAT 1919  
 QY 2949 GTTACAGAGATTAATCAACAGAACCTGATTTCTGAAAAAAGCTGACAGCCAAAC 3008  
 Db 1920 GTTACAGAGATTAATCAACAGAACCTGATTTCTGAAAAAAGCTGACAGCCAAAC 1979  
 QY 3009 TCTGCAATTTGCAATCCCACTTGTATTTGTACAGAGAGTTGATGATGATGATGAT 3068  
 Db 1980 TCTGCAATTTGCAATCCCACTTGTATTTGTACAGAGAGTTGATGATGATGATGAT 2039  
 QY 3069 GTACTATTGTGCAAG 3084  
 Db 2040 GTACTATTGTGCAAG 2055

RESULT 9  
 AAA09351  
 ID AAA09351 standard; DNA; 1302 BP.  
 AC AAA09351;  
 XX  
 DT 10-AUG-2000 (first entry)  
 DE G protein-coupled receptor protein 5 coding sequence.  
 XX  
 DE G protein-coupled receptor protein; antiparietic; antihypertensive;  
 XX GCRP-5; G protein-coupled receptor protein; antiparietic; antihypertensive;  
 XX immunosuppressive; antidiabetic; antihypertensive; antidiabetic;  
 XX antihypertensive; antidiabetic; antihypertensive; antidiabetic;  
 XX antidiabetic; antihypertensive; antidiabetic; antihypertensive;  
 XX antiparietic; immunostimulant; ss.  
 OS Homo sapiens.  
 XX  
 XX  
 Key Location/Qualifiers  
 FH 79..1035  
 FT CDS  
 FT /\*tag= a  
 XX  
 XX  
 PN WO200020590-A2.



PD 13-APR-2000.  
 XX 06-OCT-1999; 99WO-US23317.  
 PF 06-OCT-1998; 98US-0167219.  
 PR 06-OCT-1998; 98US-0172211.  
 PR 11-MAY-1999; 99US-0133585.  
 XX (INCY-) INCYTE PHARM INC.  
 PA  
 PI Tang YT, Yue H, Lal P, Bandman O, Au-cung J, Reddy R, Corley NC,  
 PI Guegler KJ, Gorgone GA, Baughn MR, Azimzal Y,  
 XX MPI; 2000-328934/28.  
 DR P-PSDB; AAY32365.  
 XX  
 PT Novel human G-protein coupled receptor proteins used in the diagnosis,  
 PT treatment and prevention of nervous system disorders,  
 PT autoimmune/inflammatory disorders, and cell proliferative disorders  
 PT such as cancer  
 PS  
 PS Claim 9; Page 78; 84p; English.  
 CC This sequence encodes human G-protein coupled receptor protein (GCRP) 5.  
 CC The GCRP polypeptides, polynucleotides, antibodies, antagonists and  
 CC agonists may be administered to human patients for the diagnosis,  
 CC treatment and prevention of nervous system disorders (e.g. epilepsy,  
 CC stroke, neoplasms, Alzheimer's disease), autoimmune or  
 CC inflammatory disorders, complications of cancer, hemodialysis and  
 CC extracorporeal circulation, and cell proliferative disorders. They are  
 CC also used to treat or prevent disorders associated with decreased or  
 CC increased expression or activity of GCRP.  
 XX  
 XX Sequence 1302 BP; 277 A; 346 C; 259 G; 420 T; 0 other;  
 Query Match 41.1%; Score 1287.8; DB 21; Length 1302;  
 Best Local Similarity 99.8%; Pred. No. 4,4e-257;  
 Matches 1300; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

DB 481 CATGCCAAGTACTTACGTTGCTGCTGCTACCAAAATGGTGTGGCTGCTGTGGCGG 540  
 QY 532 GGGGTGCTGCTGATGAGCAACCCCTTCTGCTTCAATGAGCAAGCTGCCCTTGGCGCTCC 651  
 DB 541 GGGGTGCTGCTGATGAGCAACCCCTTCTGCTTCAATGAGCAAGCTGCCCTTGGCGCTCC 600  
 QY 652 AATATCTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT 711  
 DB 601 AATATCTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT 660  
 QY 712 ATCCGGGTCAATGCTGCTATGAGCTTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 771  
 DB 661 ATCCGGGTCAATGCTGCTATGAGCTTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 720  
 QY 772 CTCTCATCT 831  
 DB 721 CTCTCATCT 780  
 QY 832 GCCCAGGCAAGGATTTGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891  
 DB 781 GCCCAGGCAAGGATTTGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 QY 892 GTACCTTTCAATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 951  
 DB 841 GTACCTTTCAATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 900  
 QY 952 CCCCTCATCTTGGCAATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1011  
 DB 901 CCCCTCATCTTGGCAATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 1012 GGAATGGAAGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAG 1071  
 DB 961 GGAATGGAAGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAG 1020  
 QY 1072 GCTTCAAGGCTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131  
 DB 1021 GCTTCAAGGCTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 QY 1132 GATTTTATGTTAATGTTTGAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGAT 1191  
 DB 1081 GATTTTATGTTAATGTTTGAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGAT 1139  
 QY 1192 TACAAGCTGATGCTTCAATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1251  
 DB 1140 TACAAGCTGATGCTTCAATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1199  
 QY 1252 TTTCTTCTTGTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1311  
 DB 1200 TTTCTTCTTGTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259  
 QY 1312 GGGTATATCTTCTTCAATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1354  
 DB 1260 GGGTATATCTTCTTCAATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1302

RESULT 10  
 AAS64174  
 ID AAS64174 standard; cDNA; 1302 BP.  
 XX  
 XX AAS64174;  
 AC  
 XX  
 DT 29-JAN-2002 (first entry)  
 XX  
 DE Human prostate cDNA sequence #593.  
 XX  
 XX Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.  
 XX  
 XX Homo sapiens.  
 XX  
 XX OS  
 XX PN W0200173032-A2.  
 XX  
 XX 04-OCT-2001.  
 PD  
 XX

PF 27-MAR-2001; 2001WO-US09919.  
 XX 27-MAR-2000; 2000US-0536857.  
 PR 09-MAY-2000; 2000US-0568100.  
 PR 12-MAY-2000; 2000US-0570737.  
 PR 13-JUN-2000; 2000US-0593793.  
 PR 27-JUN-2000; 2000US-0605783.  
 PR 10-AUG-2000; 2000US-0636215.  
 PR 29-AUG-2000; 2000US-0651236.  
 PR 06-SEP-2000; 2000US-0657279.  
 PR 02-OCT-2000; 2000US-0679426.  
 PR 10-OCT-2000; 2000US-0685166.  
 XX

(CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD,  
 PI Panzer GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skelky YAW, Hepler WT, Henderson RA;  
 DR P-PSDB; AA069951.  
 XX MPI: 2001-639232/73.

PT New human prostate-specific polypeptides and polynucleotides useful for  
 XX the diagnosis and treatment of cancer, especially prostate cancer -  
 XX Claim 1; Page 566; 579pp; English.

XX The invention relates to isolated prostate-specific  
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,  
 CC antibodies raised against the polypeptides (or antigenic epitopes  
 CC derived from them) and antigen-presenting cells expressing the  
 CC polypeptides. The antibodies are useful for detecting the presence of  
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and  
 CC the antigen-presenting cells are useful for stimulating and/or expanding  
 CC T cells specific for a tumor protein, and for inhibiting the development  
 CC of cancer especially prostate cancer. Compositions comprising the  
 CC polynucleotide and/or polypeptide are useful for stimulating an immune  
 CC response, and for treating cancer. The oligonucleotide is useful for  
 CC detecting cancer. The present sequence is a prostate specific  
 CC polynucleotide of the invention.

SQ Sequence 1302 BP; 277 A; 346 C; 259 G; 420 T; 0 other;

Query Match 41.1%; Score 1287.8; DB 22; Length 1302;  
 Best Local Similarity 99.8%; Pred. No. 4.4e-257;  
 Matches 1300; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 52 TGGACAAAGGGGGTCAACATTCCTCCATACGGTTGAGCCTTACCTGCTGCTG 111  
 DB 1 TGGACAAAGGGGGTCAACATTCCTCCATACGGTTGAGCCTTACCTGCTGCTG 60  
 QY 112 TCAAGTTCAGCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 171  
 DB 61 TCAAGTTCAGCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 120  
 QY 172 TTTATCTTAATAGGCTCTGCTTTAGAGAGGCTAGTTCTGCTGCTGCTGCT 231  
 DB 121 TTTATCTTAATAGGCTCTGCTTTAGAGAGGCTAGTTCTGCTGCTGCTGCT 180  
 QY 232 TGGTCCCTTACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 291  
 DB 181 TGGTCCCTTACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 292 GAGCAGAGCTGATGAGCAGCTATATATTTCTTTCATGCTGCTGCTGCTG 351  
 DB 241 GAGCAGAGCTGATGAGCAGCTATATATTTCTTTCATGCTGCTGCTGCTG 300  
 QY 352 CTCATTCACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411  
 DB 301 CTCATTCACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 QY 412 ATCCAGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471

DB 361 ATCCAGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
 QY 472 TCCAGTTCAGCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 531  
 DB 421 TCCAGTTCAGCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 480  
 QY 532 CAGCCAGTTCAGCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 591  
 DB 481 CAGCCAGTTCAGCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 540  
 QY 592 GAGCAGAGCTGATGAGCAGCTATATATTTCTTTCATGCTGCTGCTGCTG 651  
 DB 541 GAGCAGAGCTGATGAGCAGCTATATATTTCTTTCATGCTGCTGCTGCTG 600  
 QY 652 AATATCTTTCCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711  
 DB 601 AATATCTTTCCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
 QY 712 ATCCAGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771  
 DB 661 ATCCAGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 QY 772 CTTCATCTTCCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 831  
 DB 721 CTTCATCTTCCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
 QY 832 GCCAGGCGCAAGGATTTGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891  
 DB 781 GCCAGGCGCAAGGATTTGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
 QY 892 GTACCTTCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951  
 DB 841 GTACCTTCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 QY 952 CCGCTCATCTTGGCAATATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1011  
 DB 901 CCGCTCATCTTGGCAATATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 1012 GGAATGAAAGCAAAAGAGATTCACAGCGCATCTTCGATGCTGCTGCTG 1071  
 DB 961 GGAATGAAAGCAAAAGAGATTCACAGCGCATCTTCGATGCTGCTGCTGCTG 1020  
 QY 1072 GCTTCAGAGGCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1131  
 DB 1021 GCTTCAGAGGCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
 QY 1132 GATTTTATGTTAATGATTTGGAAGACAGTATTCAGAAAAAATTCCTTAAT 1191  
 DB 1081 GATTTTATGTTAATGATTTGGAAGACAGTATTCAGAAAAAATTCCTTAAT 1139  
 QY 1192 TACAAGTCAAGTCTTCAATATGAAAGCTGTTGGGAATTCCTTAATTTTCAAT 1251  
 DB 1140 TACAAGTCAAGTCTTCAATATGAAAGCTGTTGGGAATTCCTTAATTTTCAAT 1199  
 QY 1252 TTTCTTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1311  
 DB 1200 TTTCTTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1259  
 QY 1312 GGGTTATTAATTTTCAATTTTACATGAGTCAAGTCAAGTCAAGTCAAGT 1354  
 DB 1260 GGGTTATTAATTTTCAATTTTACATGAGTCAAGTCAAGTCAAGTCAAGT 1302

RESULT 11  
 AAH93938  
 ID AAH93938 standard; cDNA; 1302 BP.  
 XX  
 AC AAH93938;  
 XX  
 DT 04-OCT-2001 (first entry)  
 XX  
 DE P835P full length cDNA sequence.  
 XX





Db 1021 GCTTCAGAGCCCTAGGTGCTGATGATCAAACTTTTCATAGAGTCCCTTGATCA 1080  
 QY 1132 GATTATATGTTTACATTTTGAGAGACAGTATTCAGAAAAAATTTCTTATATAAAA 1191  
 Db 1081 GATTATATGTTTACATTTTGAGAGACAGTATTCAGAAAAAATTTCTTATATAAAA 1139  
 QY 1192 TACAGTGCAGTCCCTCAATATGAAACAGTGGGAAATCCCAATTTTTCATATATAT 1251  
 Db 1140 TACAGTGCAGTCCCTCAATATGAAACAGTGGGAAATCCCAATTTTTCATATATAT 1199  
 QY 1252 TTTCTTCTTGTGTTTCTGCTACATATATATATATATATATATATATATATATAT 1311  
 Db 1200 TTTCTTCTTGTGTTTCTGCTACATATATATATATATATATATATATATATATAT 1259  
 QY 1312 GGGTAT 1354  
 Db 1260 GGGTAT 1302

## RESULT 13

ACAS9982 standard; cDNA; 1302 BP.

ACAS9982:

10-JUN-2003 (first entry)

Prostate cancer therapy associated cDNA #663.

Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;

immunogen; cancer; prostate specific antigen, PSA;

prostatic acid phosphatase, PAP; prostate specific membrane antigen;

PSMA; gene; ss.

Homo sapiens.

US2002192763-A1.

19-DEC-2002.

29-JUN-2001; 2001US-0895793.

17-APR-2000; 2000US-157455P.

04-OCT-2000; 2000US-0679272.

28-MAR-2001; 2001US-0822827.

(XUJ/) XU J.

(DILL/) DILLON D C.

(MITC/) MITCHAM J L.

(HARL/) HARLOCKER S L.

(JIANG/) JIANG Y.

(KALC/) KALOS M D.

(FANG/) FANGER G R.

(RETT/) RETTER M W.

(STOL/) STOLK J A.

(DAYC/) DAY C H.

(VEDV/) VEDVICK T S.

(CART/) CARTER D.

(LISX/) LI S X.

(WANG/) WANG A.

(SKEI/) SKEIKY Y A W.

XX XE  
 DR WPI; 2003-352711/33.  
 PT New fusion protein comprising prostate-specific polypeptides, or its  
 PT immunogenic portions, useful for diagnosing, preventing and/or treating  
 PT cancer, particularly prostate cancer  
 XX Example 15; SEQ ID NO 916; 85pp; English.

CC The invention describes a fusion protein comprising at least one amino  
 CC acid sequence of immunogenic portions of any of the 3 sequences not  
 CC defined in the specification, or sequences having at least 70 or 90 %  
 CC sequence identity to any one of the 35 sequences defined in the USPTO  
 CC web site, which is encoded by any of the 4 nucleotide sequences not  
 CC defined in the specification. The fusion protein, composition and  
 CC methods are useful for diagnosing, preventing and/or treating cancer,  
 CC particularly prostate cancer. The proteins are useful as markers to  
 CC indicate the presence or absence of cancer. This sequence  
 CC represents a prostate cancer therapy associated cDNA.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from the US patent office at  
 CC seqdata.uspto.gov/sequence.html?docID=US20020192763.

Sequence 1302 BP; 277 A; 346 C; 259 G; 420 T; 0 other;

Query Match 41.1%; Score 1287.8; DB 25; Length 1302;

Best Local Similarity 99.8%; Pred. No. 4.4e-257;

Matches 1300; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 52 TGGACAAAGGGGGTACACATTTCTTCATAGAGGTGAGGCTTACCGCTGAGTGG 111  
 Db 1 TGGACAAAGGGGGTACACATTTCTTCATAGAGGTGAGGCTTACCGCTGAGTGG 60  
 QY 112 TCACAGTTGAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 171  
 Db 61 TCACAGTTGAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 QY 172 TTCATCTTATAGGCTCCCTGTTTGAAGAGGTGATGATGATGATGATGATGATGATGAT 231  
 Db 121 TTCATCTTATAGGCTCCCTGTTTGAAGAGGTGATGATGATGATGATGATGATGATGAT 180  
 QY 232 TGGTCCCTTACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 291  
 Db 181 TGGTCCCTTACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 QY 292 GAGCAAGCTGCAATGAGGCAATGATATATTTCTTTCATGATGATGATGATGATGATGATG 351  
 Db 241 GAGCAAGCTGCAATGAGGCAATGATATATTTCTTTCATGATGATGATGATGATGATGATG 300  
 QY 352 CTGATCTCAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 411  
 Db 301 CTGATCTCAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 QY 412 ATCCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 471  
 Db 361 ATCCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
 QY 472 TCCACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531  
 Db 421 TCCACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 QY 532 CAGGCAAGTACTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591  
 Db 481 CAGGCAAGTACTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 QY 592 GGGGCTGCACTATGCAACCCCTTCTGCTTTCATGATGATGATGATGATGATGATGATGAT 651  
 Db 541 GGGGCTGCACTATGCAACCCCTTCTGCTTTCATGATGATGATGATGATGATGATGATGAT 600  
 QY 652 AATATCTTCCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 711  
 Db 601 AATATCTTCCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660



QY	245	TTATTCGTGTGCTAGTAACTTGAACAATCATCTAATTGGGCGAATGGACACAGCTTC	3
Db	121	TTATTCGTGTGCTAGTAACTTGAACAATCATCTAATTGGGCGAATGGACACAGCTTC	1
QY	305	ATGAGCCCATGTATATATTTCTTTGCATGCTTTCAGGACATTGACATCTCATCTCACT	3
Db	181	ATGAGCCCATGTATATATTTCTTTGCATGCTTTCAGGACATTGACATCTCATCTCACT	2
QY	365	CATCCATGGCCAAATGCTGGGCATCTTCTGGTTCATTTCCACTACATTCAGATTGATG	4
Db	241	CATCCATGGCCAAATGCTGGGCATCTTCTGGTTCATTTCCACTACATTCAGATTGATG	3
QY	425	CTTGTCTGTACAGATTTTTCGACATCCACTGCTTATCTGACATGAAATCCACAGTGTGC	4
Db	301	CTTGTCTGTACAGATTTTTCGACATCCACTGCTTATCTGACATGAAATCCACAGTGTGC	3
QY	485	TGGCCATGGCTTTTGAACGCTATGTGGGCATCTGTACACCATGGCGCATGCCACAGTAC	5
Db	361	TGGCCATGGCTTTTGAACGCTATGTGGGCATCTGTACACCATGGCGCATGCCACAGTAC	4
QY	545	TTACGTGTGCTGTGTACCAAAATTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6
Db	421	TTACGTGTGCTGT	4
QY	605	TGGACACCCCTTCTGTCTTTCATCAAGACAGCTGCCCTTTCGCGCTCCAAATATCTTCC	6
Db	481	TGGACACCCCTTCTGTCTTTCATCAAGACAGCTGCCCTTTCGCGCTCCAAATATCTTCC	5
QY	665	ATTCCCATCGCCTTAACCAAGATGTCAAGAACTGGCGCTGTATATATTCGGGTCAATG	7
Db	541	ATTCCCATCGCCTTAACCAAGATGTCAAGAACTGGCGCTGTATATATTCGGGTCAATG	6
QY	725	TGCTCATGTGGCTTATATGTCATCAATCTCGCATTTGGCTGTGACATCACTCATCTCT	7
Db	601	TGCTCATGTGGCTTATATGTCATCAATCTCGCATTTGGCTGTGACATCACTCATCTCT	6
QY	785	TCTCATATCTGCTTATCTTAAAGACTGTGTGGGCTTGAACAAGTGAAGCCACAGCCAAAG	8
Db	661	TCTCATATCTGCTTATCTTAAAGACTGTGTGGGCTTGAACAAGTGAAGCCACAGCCAAAG	7
QY	845	CATTGGCACTTGCCTCTCATGT	9
Db	721	CATTGGCACTTGCCTCTCATGT	7
QY	905	GATGTGCATGTGCATGTCTTAAACAAGCGGCGTGACTCTCCGCTGCCCTGCATCTTGG	9
Db	781	GATGTGCATGTGCATGTCTTAAACAAGCGGCGTGACTCTCCGCTGCCCTGCATCTTGG	8
QY	965	CCAATATCATCTGCGGTGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1
Db	841	CCAATATCATCTGCGGTGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	9
QY	1025	AGGAGATTGCACAGGCAATCTTGCATCTTTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1
Db	901	AGGAGATTGCACAGGCAATCTTGCATCTTTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT	9
QY	1085	AGGTGTCA 1092	
Db	961	AGGTGTCA 968	

anabolic; cytostatic; antiviral; gene therapy; cardiomyopathy; obesity;  
KW asthenia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;  
KW asthma; Alzheimer's disease; Parkinson's disorder; Huntington's disease;  
KW infection; human immunodeficiency virus; HIV; ds.  
XX  
OS Homo sapiens.  
XX  
PN WOJ000174904-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US10241.  
XX  
PR 31-MAR-2000; 2000US-193664P.  
XX PR 05-APR-2000; 2000US-194614P.  
PR 06-APR-2000; 2000US-195063P.  
PR 06-APR-2000; 2000US-195066P.  
PR 06-APR-2000; 2000US-195067P.  
PR 06-APR-2000; 2000US-195068P.  
PR 06-APR-2000; 2000US-195069P.  
PR 06-APR-2000; 2000US-195070P.  
PR 21-JUL-2000; 2000US-219851OP.  
PR 27-JUL-2000; 2000US-221284P.  
PR 28-JUL-2000; 2000US-221325P.  
PR 11-AUG-2000; 2000US-224588P.  
PR 11-OCT-2000; 2000US-239613P.  
PR 18-JAN-2001; 2001US-262508P.  
PR 23-JAN-2001; 2001US-263433P.  
PR 23-JAN-2001; 2001US-263604P.  
PR 30-JAN-2001; 2001US-265161P.  
XX PR 29-MAR-2001; 2001US-08233172.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Majumder K, Verret CAM, Casman SJ, Wolenc AR, Spaderna SK;  
PI Padigaru M, Mishra VS, Tchiernev VT, Spytek KA, Li L;  
PI Baumgartner JC, Gusev VY;  
XX  
XX WPI: 2001-639351/73.  
DR P-PDSB; ABA44534.  
XX  
XX Claim 9; Page 57; 157pp; English.

The invention relates to nucleic acid sequences (ABA81529-ABA81552) that encode G-coupled protein-receptor related polypeptides (ABBA44523-ABBA44543). The isolated polypeptide having a sequence differing by no more than 15 % of amino acid residues from one of 22 amino acid sequences (or mature forms of the sequences), fully defined in the specification and corresponding to human G-protein coupled receptor X (GPCRX) polypeptide. The polypeptides have potential cardiac, antiarteriosclerotic, anabolic, cyostatic and antiviral activity. The therapy can be administered therapeutically, especially using gene therapy and expressing the encoding DNA in vivo, to treat or prevent GPCR-X-associated disorders, especially in humans. For example, they can be used to treat/prevent cardiomyopathy, atherosclerosis, disorders related to signal processing and metabolic pathway modulation (e.g. obesity, anorexia), diabetes, osteoporosis, Crohn's disease, multiple sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disorder, Huntington's disease), immune disorders, haematopoietic disorders, developmental diseases, neurological disorders, bacterial, fungal, protozoal and viral infections (e.g. with human immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically to determine the presence of or predisposition to a disease associated with altered levels of the polypeptide in mammals (especially humans) by detecting alterations in polypeptide expression levels relative to control samples. They are useful to identify agents binding polypeptide (e.g. cellular receptors or downstream effectors) and/or agents

modulating cellular polypeptide expression or activity, useful as antagonists and agonists in disease treatment.

Sequence 969 BP; 185 A; 282 C; 205 G; 297 T; 0 other;

Query Match	30.7%	Score 964.2;	DB 22;	Length 969;
Best Local Similarity	99.7%;	Pred. No. 4.3e-190;		
Matches 966;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	124	TTCTTCAATGAGGGGATCCCAATGGCAATGAAATCAAGGTCAATCTTCAATCTTAAT	183
Dp	1	TTCTTCAATGATGGTGAATCCCAATGGCAATGAAATCAAGGTCAATCTTCAATCTTAAT	60
QY	184	GGCTCTCCCTGCTTTAGAAAGAGCTCAATCTGCTGGCTTCCCATTTGCTCCTCTAC	243
Dp	61	GGCTCTCCCTGCTTTAGAAAGAGCTCAGTTCTGCTTGGCTTCCCATTTGCTCCTCTAC	120
QY	244	CTTATTTGCTGCTAGTAACTTTGACAAATCTTCAATTTTGCGGACTGAGCAAGCTTG	303
Dp	121	CTTATTTGCTGCTAGTAACTTTGACAAATCTTCAATTTTGCGGACTGAGCAAGCTTG	180
QY	304	CATAGGCCCAATGATATATTTCTTTGCAATGCTTCAAGSCATTTGACATCTCATCTCAAC	363
Dp	181	CATAGGCCCAATGATATATTTCTTTGCAATGCTTCAAGSCATTTGACATCTCATCTCAAC	240
QY	364	TCATTCATGGCCAAATGCTGGCAATCTTTCTGCTTCAATTCACATCAATCACTTTGAT	423
Dp	241	TCATTCATGGCCAAATGCTGGCAATCTTTCTGCTTCAATTCACATCAATCACTTTGAT	300
QY	424	GCATTGCTGTACAGATTTTTCSCAATCCACTCTTATTTGGCATGGAATCAAGATGAT	483
Dp	301	GCATTGCTGTACAGATTTTTCSCAATCCACTCTTATTTGGCATGGAATCAAGATGAT	360
QY	484	CTGGCCATGGCTTTTGAACCGCTAATGAGCAATCTGTCAACCCATGGGCAATGCAAGTA	543
Dp	361	CTGGCCATGGCTTTTGAACCGCTAATGAGCAATCTGTCAACCCATGGGCAATGCAAGTA	420
QY	544	CTTACGTTGCTGTGATGATCAACCAATTTGATGCTGTGTGGAGCGGGGGGCTGACATG	603
Dp	421	CTTACGTTGCTGTGATGATCAACCAATTTGATGCTGTGTGGAGCGGGGGGCTGACATG	480
QY	604	ATGAGCAACCCCTTCTGCTCTTCAATCAAGAGCTGACCTTCTGCGCTTCAATCTCTTCC	663
Dp	481	ATGAGCAACCCCTTCTGCTCTTCAATCAAGAGCTGACCTTCTGCGCTCTCAATCTCTTCC	540
QY	664	CATTCTCAATGCTCAACCAAGATGATCAATGAAAGCTGGGCTGTGATGATATCCGGGCTCAT	723
Dp	541	CATTCTCAATGCTCAACCAAGATGATCAATGAAAGCTGGGCTGTGATGATATCCGGGCTCAT	600
QY	724	GTCGTATATGACTTATGCTCATCATCTCCGCAATTTGAGCTGAGACTCACTTCTCATCTCC	783
Dp	601	GTCGTATATGACTTATGCTCATCATCTCCGCAATTTGAGCTGAGACTCACTTCTCATCTCC	660
QY	784	TTCTCATATCTGCTTATCTTTAAGCTGTGTTGGGCTTTGACAGTGAAGCCAGGCAAG	843
Dp	661	TTCTCATATCTGCTTATCTTTAAGCTGTGTTGGGCTTTGACAGTGAAGCCAGGCAAG	720
QY	844	GCATTTTGGCACTTGGGCTCTCATGTGATGCTGTGCTATTTCAATTTCAATGATCACTTCAT	903
Dp	721	GCATTTTGGCACTTGGGCTCTCATGTGATGCTGTGCTATTTCAATTTCAATGATCACTTCAT	780
QY	904	GGATTTGTCATGGTGCATGCTTTTAGCAAGGCGGTGATCTCTCCGCTGCCCTCATCTTG	963
Dp	781	GGATTTGTCATGGTGCATGCTTTTAGCAAGGCGGTGATCTCTCCGCTGCCCTCATCTTG	840
QY	964	GGCAATATCTATCTGCTGGTCTCTCTGCTGCTCAACCAATTTGATATGAGATGAAGACA	1022
Dp	841	GGCAATATCTATCTGCTGGTCTCTCTGCTGCTCAACCAATTTGATATGAGATGAAGACA	900
QY	1024	AAGGATTTGCAAGGCGCATCTCTTGACTTTTTCATATGAGCCACACAGCTTGAAGGCC	1083
Dp	901	AAGGATTTGCAAGGCGCATCTCTTGACTTTTTCATATGAGCCACACAGCTTGAAGGCC	960

QY	1084	TAGGTGTCA	1092
Db	961	TAGGTGTCA	969

```
Search completed: February 9, 2004, 19:33:26
Job time : 808 secs
```





Tue Feb 10 06:04:26 2004

us-10-017-066a-2.rat

Page 2

Db 244 IGVLAAYVPLIGSVHFRGNSLHPFIVRVMGDIYLLPVPINPIYGAKTKOIRTRVLA MFKIS 303  
QY 306 RLPHVA 311  
Db 304 AMFKIS 309

RESULT 2  
US-08-465-980-2  
Sequence 2, Application US/08465980  
Patent No. 5756309  
GENERAL INFORMATION:  
APPLICANT: Soppet, Daniel R.  
APPLICANT: Li, Yi  
APPLICANT: Rosen, Craig A.  
APPLICANT: Ruben, Steven M.  
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRJ70  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
ADDRESSEE: STUART & OLSTEIN  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,980  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-446  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 320 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-465-980-2

Query Match 59.6%; Score 964.5; DB 1; Length 320;  
Best Local Similarity 59.4%; Pred. No. 2.6e-80;  
Matches 177; Conservative 54; Mismatches 66; Indels 1; Gaps 1;  
QY 15 ILGLGLEBAQWLAFLPCLSLYLIVLGNLTITTYVTRTSHLHEBYVFLCMLSGDIL 74  
DB 12 VILGIGLEBAHFWGFPPLSMYVAMCNCIVFVTRERSLHAPMYFLCMLAIDLA 71  
QY 75 ISTDMPKMLAIFWPNSTTIQFDACLLQFAHISLGMESTVLLMAFDRYVAICPLRH 134  
DB 72 LSTSTMPKILALFWPDSREISIRACLTQMFPIHALSAISTILLMAFDRYVAICPLRH 131  
QY 135 ATVTLPRVTKIGVAAVVRGAALMAPLPVFIKQLPFCRSNITLSHSYCLHODVWKACDDI 194  
DB 132 AAVLNTVTVAQIGIVAVRGSLEFFPLPIKRLAFCHSNVLSHSYCVHODVWKLAAYDT 191  
QY 195 RVNVVYGLIYVISAIGDLSLISFSYLLIKTVLGL-TREAQAKAFGTCVSHVCAVFIY 253  
DB 192 LRVNVYGLTALILVWGDVWFISLSTFLIRIVLQLPKSRKAKAGTCVSHIGVLAIFY 251  
QY 254 VPFILGSMVHRFSKRSDPLPVILANITYLLVPPVNLPIYGVKTEIKRILRLPHVA 311

Db 252 VPLIGLSVHFRGNSLHPFIVRVMGDIYLLPVPINPIYGAKTKOIRTRVLA MFKIS 309  
RESULT 3  
US-09-053-303-2  
Sequence 2, Application US/09053303  
Patent No. 5948890  
GENERAL INFORMATION:  
APPLICANT: Soppet, Daniel R.  
APPLICANT: Li, Yi  
APPLICANT: Rosen, Craig A.  
APPLICANT: Ruben, Steven M.  
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRJ70  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
ADDRESSEE: STUART & OLSTEIN  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/053,303  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/465,980  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-446  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 320 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-053-303-2

Query Match 59.6%; Score 964.5; DB 2; Length 320;  
Best Local Similarity 59.4%; Pred. No. 2.6e-80;  
Matches 177; Conservative 54; Mismatches 66; Indels 1; Gaps 1;

QY 15 ILGLGLEBAQWLAFLPCLSLYLIVLGNLTITTYVTRTSHLHEBYVFLCMLSGDIL 74  
DB 12 VILGIGLEBAHFWGFPPLSMYVAMCNCIVFVTRERSLHAPMYFLCMLAIDLA 71  
QY 75 ISTDMPKMLAIFWPNSTTIQFDACLLQFAHISLGMESTVLLMAFDRYVAICPLRH 134  
DB 72 LSTSTMPKILALFWPDSREISIRACLTQMFPIHALSAISTILLMAFDRYVAICPLRH 131  
QY 135 ATVTLPRVTKIGVAAVVRGAALMAPLPVFIKQLPFCRSNITLSHSYCLHODVWKACDDI 194  
DB 132 AAVLNTVTVAQIGIVAVRGSLEFFPLPIKRLAFCHSNVLSHSYCVHODVWKLAAYDT 191  
QY 195 RVNVVYGLIYVISAIGDLSLISFSYLLIKTVLGL-TREAQAKAFGTCVSHVCAVFIY 253  
DB 192 LRVNVYGLTALILVWGDVWFISLSTFLIRIVLQLPKSRKAKAGTCVSHIGVLAIFY 251  
QY 254 VPFILGSMVHRFSKRSDPLPVILANITYLLVPPVNLPIYGVKTEIKRILRLPHVA 311  
DB 252 VPLIGLSVHFRGNSLHPFIVRVMGDIYLLPVPINPIYGAKTKOIRTRVLA MFKIS 309

RESULT 4  
US-09-339-115-2  
Sequence 2, Application US/0939115  
Patent No. 6372891  
GENERAL INFORMATION:  
APPLICANT: Soppet, Daniel R.  
APPLICANT: Li, Yi  
APPLICANT: Rosen, Craig A.  
APPLICANT: Ruben, Steven M.  
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
ADDRESSEE: STUART & OLSTEIN  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/339,115  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/053,303  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-446  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1744  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 320 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-339-115-2

Query Match 59.6%; Score 964.5; DB 4; Length 320;  
Best Local Similarity 59.4%; Pred. No. 2.6e-80;  
Matches 177; Conservative 54; Mismatches 66; Indels 1; Gaps 1;  
QY 15 ILIGPGLAEQFWLAFPLCSLYLAVGNLTIIYVREHSLHEPMYFLCMLSGIDIL 74  
DB 12 VLIIGPGLAEKAFWGFPLSMYVAMGNCIVFIVTERSLHAPWYFLCMLAIDLA 71  
QY 75 ISTSSPKMLAIFWENSTTIQFDACLOTFATHSLSGMESTVLLAMAFDRVVAICHPLRH 134  
DB 72 LSTSTMPKILALFWFDSREISIEACLTQMFIFHALSAESTITLLAMAFDRVVAICHPLRH 131  
QY 135 ATVLTLPRTYKIGVAAVVGALMAPLPVFIKQLPFCRSNLTSHSYCLHODVMKACDDI 194  
DB 132 AAVLNTVTQAQIGIAVAVRGSLPFPLPLIKRLAFCHSNVLSHSYCVHODVMKLAAYDT 191  
QY 195 RVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGL-TREAQAKAFGTCVSHVCAVFIY 253  
DB 192 LPNVVYGLTALLVGVDMFSLSYFLIRTVQLPSKSRKAKAFGTCVSHIGVLAIFY 251  
QY 254 VPIGLSVVHRFGNSLHPIVRVVMDIYLLPVPINPIIYGAKTKQIRTVLAMFKIS 311  
DB 252 VPIGLSVVHRFGNSLHPIVRVVMDIYLLPVPINPIIYGAKTKQIRTVLAMFKIS 309

RESULT 5

PCT-US95-07093-2  
Sequence 2, Application PC/TUS9507093  
GENERAL INFORMATION:  
APPLICANT: Soppet, Daniel R.  
APPLICANT: Li, Yi  
APPLICANT: Rosen, Craig A.  
APPLICANT: Ruben, Steven M.  
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
ADDRESSEE: STUART & OLSTEIN  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07093  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/465,980  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-446  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 320 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-07093-2

Query Match 59.6%; Score 964.5; DB 5; Length 320;  
Best Local Similarity 59.4%; Pred. No. 2.6e-80;  
Matches 177; Conservative 54; Mismatches 66; Indels 1; Gaps 1;  
QY 15 ILIGPGLAEQFWLAFPLCSLYLAVGNLTIIYVREHSLHEPMYFLCMLSGIDIL 74  
DB 12 VLIIGPGLAEKAFWGFPLSMYVAMGNCIVFIVTERSLHAPWYFLCMLAIDLA 71  
QY 75 ISTSSPKMLAIFWENSTTIQFDACLOTFATHSLSGMESTVLLAMAFDRVVAICHPLRH 134  
DB 72 LSTSTMPKILALFWFDSREISIEACLTQMFIFHALSAESTITLLAMAFDRVVAICHPLRH 131  
QY 135 ATVLTLPRTYKIGVAAVVGALMAPLPVFIKQLPFCRSNLTSHSYCLHODVMKACDDI 194  
DB 132 AAVLNTVTQAQIGIAVAVRGSLPFPLPLIKRLAFCHSNVLSHSYCVHODVMKLAAYDT 191  
QY 195 RVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGL-TREAQAKAFGTCVSHVCAVFIY 253  
DB 192 LPNVVYGLTALLVGVDMFSLSYFLIRTVQLPSKSRKAKAFGTCVSHIGVLAIFY 251  
QY 254 VPIGLSVVHRFGNSLHPIVRVVMDIYLLPVPINPIIYGAKTKQIRTVLAMFKIS 311  
DB 252 VPIGLSVVHRFGNSLHPIVRVVMDIYLLPVPINPIIYGAKTKQIRTVLAMFKIS 309

RESULT 6  
US-08-988-876-7  
Sequence 7, Application US/08988876  
Patent No. 6063596

GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED  
TITLE OF INVENTION: WITH IMMUNE RESPONSE  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,876  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0441 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 314 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 32086  
US-08-988-876-7

Query Match 26.4%; Score 427.5; DB 3; Length 314;  
Best Local Similarity 34.5%; Pred. No. 2.2e-31;  
Matches 108; Conservative 64; Mismatches 120; Indels 21; Gaps 8;

7 NEESSATYFLILGPGLEBAQFWLAFPLGSLYLAVGNITTIYVTERSHHEPMTIFLC 66  
5 NQISIDFLLHGFIOPEQONTCYALFLAMVLTLLGLNLIIVLRDUSHLHTPEYFLS 64  
67 MNSGIDILISTSMPPKLAIFWNSSTTIQFDACLLQIFAIHSLSGMESTVLLAMAFDRYV 126  
65 NLSFSDLCFSSVTLIPFLQNMONOPSIYADCLQWFFLLFGDLSEFLVAMAYDRYV 124  
127 ALCHPLRHATVLTLPRTYKIGVA---AVVRGALAAAPLPVFIKQLPFCNSNLSHSCYCLH 183  
125 ALCFPL-RHTAISMPLCLALVALSVLTFHAMHTL--LMARLCFQADVVIPIHFFCDM 181  
184 QVWKACDIDIRVN---VVGGLVVISAGLSDLLISFSYLLIKTVLGL-TREAQAK 237  
182 SALLKIAFDRTVNEVPIRMGLIIVT---PFLILIGSVARIVSSILKVPSSKXICK 236  
238 AAGTCASHCAFIFFVPIFGISMHRPSKRDSPL-EYILANITLLVPLVPLNPIVGVK 296  
237 ASSTGSHLSVSLFYGTIVGLYC--SSANSSTLKOTVAMMYTVTPMLNPIYSIR 293  
297 TREIRORILRFLPH 309  
294 NRDMKGAISRVIH 306

RESULT 7  
US-08-988-876-5  
Sequence 5, Application US/08988876  
Patent No. 6063596  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED  
TITLE OF INVENTION: WITH IMMUNE RESPONSE  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,876  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0441 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 309 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1314667  
US-08-988-876-5

Query Match 24.4%; Score 394.5; DB 3; Length 309;  
Best Local Similarity 29.0%; Pred. No. 2.1e-28;  
Matches 87; Conservative 76; Mismatches 126; Indels 11; Gaps 5;

7 NEESSATYFLILGPGLEBAQFWLAFPLGSLYLAVGNITTIYVTERSHHEPMTIFLC 66  
5 NDRIRPFLILGSEBPKQPLFLGLFSLMVLTVTLIGNLLILAVSSDHLHTPMYFLA 64  
67 MNSGIDILISTSMPPKLAIFWNSSTTIQFDACLLQIFAIHSLSGMESTVLLAMAFDRYV 126  
65 NLSFVDFICFTCTIPGLVNIQTKVITYESCIILQMYFFELFAGIDNFLTMAYDRYN 124  
127 ALCHPLRHATVLTLPRTYKIGVA---AVVRGALAAAPLPVFIKQLPFCNSNLSHSCYCLH 186  
125 ALCFPLHATVIMKPOLCSLLVSWTMSALHSLQTLMLRLSFTCHFOIPHFFCELNM 184  
187 MKLACDIDIRVN---VVGGLVVISAGLSDLLISFSYLLIKTVLGL-TREAQAKAFGTVC 243  
185 IQLAGSDTFLNMNMVLYFAAILGVAPLVGL--YSPFKVSSIRGISASHSKYKAFSTCA 242  
244 SHVCAVFIFFVPIG--LSMHRFSKRKDSPLVILANITLLVPLVPLNPIVGVKTKER 301

Db 243 SHLSVSLRYCTSLGYSLSAAPSQTHSS-----VASVYTVVTPMLNPFYLSLNKDK 298

RESULT 8  
US-08-988-876-6  
Sequence 6, Application US/08988876  
Patent No. 6063596  
GENERAL INFORMATION:  
APPLICANT: Lai, Preeti  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED  
TITLE OF INVENTION: WITH IMMUNE RESPONSE  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,876  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0441 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 333 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 205814  
US-08-988-876-6

Query Match 23.7%; Score 383; DB 3; Length 333;  
Best Local Similarity 28.3%; Pred. No. 2,6e-27;  
Matches 94; Conservative 73; Mismatches 133; Indels 32; Gaps 9;

QY 5 NNESSATYFILILGPGLEAOFMLAPLCSLYLAIVAGNTIYIVTESHLEPMYIF 64  
Db 3 SSNRTRVSEFFILGVEKNDQPLIYGLSLSVLTIVIGNISIIVAIISDCLHTPMYFF 62

QY 65 LCMISGIDILISTSSMPMLAIFWNSSTTIQPDACLIQIFAIHSISGMSVLLAMAFDR 124  
Db 63 LSNLSFVDICISTVPMKLVVIGQNNVITYAGCITQIYFFLFEVDNDFLITIMAYDR 122

QY 125 YVAICHPLRHATVLLP-----RYTKIGVAAYVRGAALAPVPFKQLPFRSNILSHSY 180  
Db 123 YVAICHPMHYIVIMYKLCGLFVLSVTSVLAH--ALPQSL--WMLALPCTHLEIPHYF 178

QY 181 CLHODVMKLCADDIVN--VYGLIIVISAIGLDSLISFSYLLIKTVGLTR-EAQAQ 237  
Db 179 CEPNOVIQLTGDAFLNDLVIVFTLVLATVPLAG--IFSYFNTVSSICAISSVHKYK 236

QY 238 AFGTGVSHVCAVFIFVYVPIGLSMVHRFSKRDSPLFVILANI-YLLVPEVLPNIYGVK 296  
Db 237 AFGTCASHLSVSLFYCTGLG---VYLSAANNSSQASATASVMTYVTPVNPFIYSLR 293

QY 297 TKRIHQ-----RIIRLFHVATH 313  
Db 294 NKDVSVLKKTLCBEVYRSPPSILHFFLVLCR 325

RESULT 9  
US-09-465-901-48  
Sequence 48, Application US/09465901  
Patent No. 6492143  
GENERAL INFORMATION:  
APPLICANT: Reed, Randall  
APPLICANT: Kauwurst, Dietmar  
APPLICANT: Yau, King-Wai  
TITLE OF INVENTION: Olfactory Receptor Expression Libraries  
TITLE OF INVENTION: ad Methods of Making and Using Them  
FILE REFERENCE: 00107.00105  
CURRENT APPLICATION NUMBER: US/09/465,901  
CURRENT FILING DATE: 2001-12-17  
PRIOR APPLICATION NUMBER: 60/112,605  
PRIOR FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 48  
LENGTH: 313  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
OTHER INFORMATION: PCR primer  
US-09-465-901-48

Query Match 23.1%; Score 374; DB 4; Length 313;  
Best Local Similarity 32.6%; Pred. No. 1,6e-26;  
Matches 98; Conservative 56; Mismatches 137; Indels 10; Gaps 6;

QY 7 NNESSATYFILILGPGLEAOFMLAPLCSLYLAIVAGNTIYIVTESHLEPMYIFC 66  
Db 3 NSTTVEFFILGSDACEQVLIFFGLTYFLLIGNELIIFITLVDRRLVPMYFFLR 62

QY 67 MMSGIDILISTSSMPMLAIFWNSSTTIQPDACLIQIFAIHSISGMSVLLAMAFDRY 126  
Db 63 NFAMLEIWFPSVIFPMLTNIITGHKTIISLGCFQAFYFFLGTEFFLLAMSPDRYV 122

QY 127 AICHPLRHATVLLPRTVKIGVAAYVRGAALAPLPVPFKQLPFRSNILSHSYCLHODV 186  
Db 123 AICNPLRATIMKRKVCQVIFCSWMSGLLIIIVBSIYFQDPFCPPNIINHFCCNPF 182

QY 187 MKLACDDIRNVVYGLIIVTISAIGLDSLISFS-YLLIKTVGL-TRDAQAAFGTCVS 244  
Db 183 MELICADTSLVEPLGVIANFSI-LGTLAVTATCYGHIIYTLIHPSAKERKKAFCSS 241

QY 245 HVCAGFIYVYFGLSMVHRFSR---RDSPLVILANIYLLVPPVLPNIYGVTKRIK 301  
Db 242 HIIIVSLFQSGCI--FMYVRSKGNGGEDHMKVALLN--TVVTPLNFFIYTLRKQVK 297

QY 302 Q 302  
Db 298 Q 298

RESULT 10  
US-08-467-948a-2  
Sequence 2, Application US/08467948A  
Patent No. 5998164  
GENERAL INFORMATION:  
APPLICANT: LI, YI  
APPLICANT: CAO, LIANG  
APPLICANT: NI, JUAN  
APPLICANT: GENTZ, REINER

APPLICANT: BULT, CAROL J.  
APPLICANT: SUTTON III, GRANGER G.  
APPLICANT: ROSEN, CRAIG A.  
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,948A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04079  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: STEEFE, ERIC K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-467-948A-2

Query Match 22.7%; Score 367; DB 2; Length 296;  
Best Local Similarity 32.2%; Pred. No. 6.5e-26;  
Matches 94; Conservative 61; Mismatches 121; Indels 16; Gaps 7;  
QY 7 NESAATYFIILGL---PGLAEAQFWLAFPLCSL-YLIAVGNLTIIYIVRTHSHLHPY 62  
DB 8 NQWVTEFLGLGFLGRIO---MLLFGLSLFYVFTLLGNTLIGLISLDSRLHTPY 63  
QY 63 IFCLMSGIDILISTSMFKMLAFWNSSTTIOFDACLOIFAHSLSGNESTVLLAMAF 122  
DB 64 FFLSHLAVNIAVACNTVPQMLVNLHPAKPISFAGCMTLDFLSPAHTECLLVMSY 123  
QY 123 DRYVAICHPLRHAIVTLPRVTKIGVAAVRGAALMAPLVFVKQDPFCRSNLSHSCYL 182  
DB 124 DRYVAICHPLRHAIVTLPRVTKIGVAAVRGAALMAPLVFVKQDPFCRSNLSHSCYL 183  
QY 183 HQDWKLAACDDIRVN--VYVGLVVISAGLSDLLISFSYLL--ILKTVGLTREAOAKA 238  
DB 184 ILTVLRACADTWMNQVVFACMFLVGPLCLVLSVSHLGGILRIQSG--EGRRKA 240  
QY 239 FGLTVSHVCAVFIYVVFIFGLSWVHRFSKRDSPLPILANIYLVPPVLP 290  
DB 241 FSTCSSHLGVGLFFGSAIYVMAPK-SRHPBEOQKVLFILOFLSTPMKP 291

RESULT 11  
US-08-467-947A-2  
Sequence 2, Application US/08467947A  
Patent No. 6090575  
GENERAL INFORMATION:  
APPLICANT: LI, YI  
APPLICANT: CAO, LIANG  
APPLICANT: NI, JIAN

APPLICANT: GENTZ, REINER  
APPLICANT: BULT, CAROL J.  
APPLICANT: SUTTON III, GRANGER G.  
APPLICANT: ROSEN, CRAIG A.  
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,947A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04079  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: STEEFE, ERIC K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-467-947A-2

Query Match 22.7%; Score 367; DB 3; Length 296;  
Best Local Similarity 32.2%; Pred. No. 6.5e-26;  
Matches 94; Conservative 61; Mismatches 121; Indels 16; Gaps 7;  
QY 7 NESAATYFIILGL---PGLAEAQFWLAFPLCSL-YLIAVGNLTIIYIVRTHSHLHPY 62  
DB 8 NQWVTEFLGLGFLGRIO---MLLFGLSLFYVFTLLGNTLIGLISLDSRLHTPY 63  
QY 63 IFCLMSGIDILISTSMFKMLAFWNSSTTIOFDACLOIFAHSLSGNESTVLLAMAF 122  
DB 64 FFLSHLAVNIAVACNTVPQMLVNLHPAKPISFAGCMTLDFLSPAHTECLLVMSY 123  
QY 123 DRYVAICHPLRHAIVTLPRVTKIGVAAVRGAALMAPLVFVKQDPFCRSNLSHSCYL 182  
DB 124 DRYVAICHPLRHAIVTLPRVTKIGVAAVRGAALMAPLVFVKQDPFCRSNLSHSCYL 183  
QY 183 HQDWKLAACDDIRVN--VYVGLVVISAGLSDLLISFSYLL--ILKTVGLTREAOAKA 238  
DB 184 ILTVLRACADTWMNQVVFACMFLVGPLCLVLSVSHLGGILRIQSG--EGRRKA 240  
QY 239 FGLTVSHVCAVFIYVVFIFGLSWVHRFSKRDSPLPILANIYLVPPVLP 290  
DB 241 FSTCSSHLGVGLFFGSAIYVMAPK-SRHPBEOQKVLFILOFLSTPMKP 291

RESULT 12  
US-08-465-980-3  
Sequence 3, Application US/08465980  
Patent No. 5756309  
GENERAL INFORMATION:  
APPLICANT: Soppet, Daniel R.  
APPLICANT: LI, YI

APPLICANT: Rosen, Craig A.  
ADDRESSEE: Ruben, Steven M.  
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAL70  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
ADDRESSEE: STUART & OLSTEIN  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,980  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-446  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1744  
TELEFAX: 201-994-1700  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-465-980-3

Query Match 22.1%; Score 357.5; DB 1; Length 247;  
Best Local Similarity 36.5%; Pred. No. 3.9e-25;  
Matches 93; Conservative 46; Mismatches 99; Indels 17; Gaps 6;

14 FILIGPGLEAQLAPFLCGLYLAIVGNLTIIYVTSLSHEPMTIFCLMSGDI 73  
1 FILIGPGLEAQLAPFLCGLYLAIVGNLTIIYVTSLSHEPMTIFCLMSGDI 60  
74 LISTSMPTALIFWNSSTTIOFDACLIQIFAIHSLGMESTVLAMADRYVAICPLR 133  
61 CFSSTVITPRLQNMQNDSPYADCLTQMYFLLFGDLESFLVAMADRYVAICPLR 119  
134 HATVLTLPVTKIGVA--AVVRGAALMAPLVFIKQLPFCRSNLSHSYCLHODVMKLA 190  
120 HTTAISMPLCLALVALSWLTFTHAMHTL--LVARLCFCADNVIPHFCDMSALLKLA 177  
191 CDDIRVN-----VYGLIIVISAIGDSLISFSYLLIKTVLGL-TREAQAKAFGTGVS 244  
178 FSDTRNEMVIFIMGILIVT-----PRLILGSYARIVSSILKVPSSGIGKASTGCS 232  
245 HVCANFIYVPPIGL 259  
233 HLSVSLFYGVITGL 247

RESULT 13  
US-09-053-303-3  
Sequence 3, Application US/09053303  
Patent No. 594880  
GENERAL INFORMATION:  
APPLICANT: Soppet, Daniel R.  
APPLICANT: Li, Yi  
APPLICANT: Rosen, Craig A.  
APPLICANT: Ruben, Steven M.  
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAL70  
NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
ADDRESSEE: STUART & OLSTEIN  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/053,303  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/465,980  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-446  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-053-303-3

Query Match 22.1%; Score 357.5; DB 2; Length 247;  
Best Local Similarity 36.5%; Pred. No. 3.9e-25;  
Matches 93; Conservative 46; Mismatches 99; Indels 17; Gaps 6;

14 FILIGPGLEAQLAPFLCGLYLAIVGNLTIIYVTSLSHEPMTIFCLMSGDI 73  
1 FILIGPGLEAQLAPFLCGLYLAIVGNLTIIYVTSLSHEPMTIFCLMSGDI 60  
74 LISTSMPTALIFWNSSTTIOFDACLIQIFAIHSLGMESTVLAMADRYVAICPLR 133  
61 CFSSTVITPRLQNMQNDSPYADCLTQMYFLLFGDLESFLVAMADRYVAICPLR 119  
134 HATVLTLPVTKIGVA--AVVRGAALMAPLVFIKQLPFCRSNLSHSYCLHODVMKLA 190  
120 HTTAISMPLCLALVALSWLTFTHAMHTL--LVARLCFCADNVIPHFCDMSALLKLA 177  
191 CDDIRVN-----VYGLIIVISAIGDSLISFSYLLIKTVLGL-TREAQAKAFGTGVS 244  
178 FSDTRNEMVIFIMGILIVT-----PRLILGSYARIVSSILKVPSSGIGKASTGCS 232  
245 HVCANFIYVPPIGL 259  
233 HLSVSLFYGVITGL 247

RESULT 14  
US-09-339-115-3  
Sequence 3, Application US/09339115  
Patent No. 6372851  
GENERAL INFORMATION:  
APPLICANT: Soppet, Daniel R.  
APPLICANT: Li, Yi  
APPLICANT: Rosen, Craig A.  
APPLICANT: Ruben, Steven M.  
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAL70  
CORRESPONDENCE ADDRESS: 8

Tue Feb 10 06:04:26 2004

us-10-017-066a-2.ra1

Page 8

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
ADDRESSEE: STUART & OLSTEIN  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/339,115  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/053,303  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-446  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1744  
TELEFAX: 201-994-1700  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-339-115-3

Query Match 22.1%; Score 357.5; DB 4; Length 247;  
Best Local Similarity 36.5%; Pred. No. 3.9e-25;  
Matches 93; Conservative 46; Mismatches 99; Indels 17; Gaps 6;  
QY 14 FILIGLPLEEACQFWLAFLPLCSLYLAVGNLTITIVVREHSHHEPMYIFLCYLSGIDI 73  
DB 1 FILIGLPIQPEQQLCYALFLAWLTLLGNLLIIVIRDSHHTMTMFLSNLSFSDL 60  
QY 74 LISTSSMPKMLAFWENSTTIOFDACILQIFAIHSLSGMESTVLLAMAFDRYVAICHPLR 133  
DB 61 CFSSVTIPKLLQNNQNDPSIPYADCLTQWYFLLFGDLESFLVAMAYDRYVAICFPL- 119  
QY 134 HATVLTLPRTKIGVA--AVRGAALMAPLPIKOLPFCRSNIISHSYCLHODWKLA 190  
DB 120 HTYIMSPMLCALVALSWLITTFHAMLTL--LMARLCFCADNVIPHFCDMSALIKLA 177  
QY 191 CDDIRN----VYGLIVITSAIGDSLISFSYLLIKTVGL-TREAQARAFTCVS 244  
DB 178 FSDTRVEMWIFIMGILIVI-----PFLILGSYARIVSILKVPSSKGIKRAFSTGGS 232  
QY 245 HVCAPFIYVPIGL 259  
DB 233 HLSVSLFYGTIVGL 247

STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07093  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/465,980  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-446  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1744  
TELEFAX: 201-994-1700  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-07093-3

Query Match 22.1%; Score 357.5; DB 5; Length 247;  
Best Local Similarity 36.5%; Pred. No. 3.9e-25;  
Matches 93; Conservative 46; Mismatches 99; Indels 17; Gaps 6;  
QY 14 FILIGLPLEEACQFWLAFLPLCSLYLAVGNLTITIVVREHSHHEPMYIFLCYLSGIDI 73  
DB 1 FILIGLPIQPEQQLCYALFLAWLTLLGNLLIIVIRDSHHTMTMFLSNLSFSDL 60  
QY 74 LISTSSMPKMLAFWENSTTIOFDACILQIFAIHSLSGMESTVLLAMAFDRYVAICHPLR 133  
DB 61 CFSSVTIPKLLQNNQNDPSIPYADCLTQWYFLLFGDLESFLVAMAYDRYVAICFPL- 119  
QY 134 HATVLTLPRTKIGVA--AVRGAALMAPLPIKOLPFCRSNIISHSYCLHODWKLA 190  
DB 120 HTYIMSPMLCALVALSWLITTFHAMLTL--LMARLCFCADNVIPHFCDMSALIKLA 177  
QY 191 CDDIRN----VYGLIVITSAIGDSLISFSYLLIKTVGL-TREAQARAFTCVS 244  
DB 178 FSDTRVEMWIFIMGILIVI-----PFLILGSYARIVSILKVPSSKGIKRAFSTGGS 232  
QY 245 HVCAPFIYVPIGL 259  
DB 233 HLSVSLFYGTIVGL 247

Search completed: February 9, 2004, 16:26:17  
Job time : 22 sec



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 16:25:17 ; Search time 33 Seconds

(without alignments)  
2011.339 Million cell updates/sec

Title: US-10-017-066a-2

Sequence: 1 MVDPNGNSATYFILGLP.....KEIRKRLFLHVATASRP 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCF\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/FCFUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1617	100.0	317	US-10-205-823-443	Sequence 443, App
2	1617	100.0	318	US-10-001-469-2880	Sequence 2880, App
3	1614	99.8	317	US-10-295-027-873	Sequence 873, App
4	1614	99.8	317	US-10-293-798-746	Sequence 746, App
5	1614	99.8	318	US-09-759-143-920	Sequence 920, App
6	1614	99.8	318	US-09-780-669-920	Sequence 920, App
7	1614	99.8	318	US-09-822-827-920	Sequence 920, App
8	1614	99.8	318	US-09-886-055-91	Sequence 91, App1
9	1614	99.8	318	US-09-895-793-920	Sequence 920, App
10	1614	99.8	318	US-09-895-814-920	Sequence 920, App
11	1614	99.8	318	US-09-966-459A-2	Sequence 2, App1
12	1614	99.8	318	US-09-804-291-91	Sequence 91, App1
13	1614	99.8	318	US-10-017-161-868	Sequence 868, App
14	1614	99.8	318	US-10-144-678A-920	Sequence 920, App
15	1614	99.8	318	US-10-294-025-920	Sequence 920, App

16	1614	99.8	318	US-10-044-643-26	Sequence 26, App1
17	1614	99.8	318	US-10-012-896-920	Sequence 920, App
18	1614	99.8	318	US-10-225-667A-434	Sequence 434, App
19	1607	99.4	318	US-10-044-643-28	Sequence 28, App1
20	1607	99.4	318	US-10-044-643-30	Sequence 30, App1
21	1542	95.4	303	US-10-001-469-2887	Sequence 2887, App
22	1534	94.9	302	US-10-001-469-2882	Sequence 2882, App
23	1515	93.7	298	US-10-001-469-2883	Sequence 2883, App
24	1515	93.7	298	US-10-001-469-2885	Sequence 2885, App
25	1262	78.0	249	US-10-114-669-8894	Sequence 8894, App
26	985	60.9	320	US-10-005-041A-42	Sequence 42, App1
27	979	60.5	318	US-10-024-399-18	Sequence 18, App1
28	979	60.5	320	US-09-759-143-527	Sequence 527, App1
29	979	60.5	320	US-09-730-018-7	Sequence 7, App1
30	979	60.5	320	US-09-780-669-527	Sequence 527, App
31	979	60.5	320	US-09-822-827-527	Sequence 527, App
32	979	60.5	320	US-09-886-055-83	Sequence 83, App1
33	979	60.5	320	US-09-868-033C-4	Sequence 4, App1
34	979	60.5	320	US-09-895-793-527	Sequence 527, App
35	979	60.5	320	US-09-895-814-527	Sequence 527, App
36	979	60.5	320	US-09-804-291-83	Sequence 83, App1
37	979	60.5	320	US-10-144-678A-527	Sequence 527, App
38	979	60.5	320	US-10-294-025-527	Sequence 527, App
39	979	60.5	320	US-10-044-643-65	Sequence 65, App1
40	979	60.5	320	US-10-431-842-7	Sequence 7, App1
41	979	60.5	320	US-10-025-806-34	Sequence 34, App1
42	979	60.5	320	US-10-387-629-200	Sequence 200, App
43	979	60.5	320	US-10-005-041A-41	Sequence 41, App1
44	979	60.5	320	US-09-581-566A-55	Sequence 55, App1
45	979	60.5	320	US-10-012-896-527	Sequence 527, App

#### ALIGNMENTS

RESULT 1  
US-10-205-823-443  
Sequence 443, Application US/10205823  
Publication No. US20030108963A1  
GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Monahan, John E.  
APPLICANT: Endege, Wilson O.  
APPLICANT: Gannavarapu, Manjula  
APPLICANT: Gorbacheva, Bella  
APPLICANT: Hoersch, Sebastian  
APPLICANT: Kanakkar, Shubhangi  
APPLICANT: Monsey, Angela M.  
APPLICANT: Glatz, Karen  
APPLICANT: Zhao, Xumei  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
THERAPY OF PROSTATE CANCER  
FILE REFERENCE: MRI-044  
CURRENT APPLICATION NUMBER: US/10/205,823  
PRIOR FILING DATE: 2002-07-25  
PRIOR APPLICATION NUMBER: 60/307,982  
PRIOR FILING DATE: 2001-07-25  
PRIOR APPLICATION NUMBER: 60/314,356  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/325,020  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: 60/341,746  
PRIOR FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: 60/362,158  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 443  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens

US-10-205-823-443

Query Match 100.0%; Score 1617; DB 15; Length 317;

Best Local Similarity 100.0%; Pred. No. 5,9e-150; Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVDNNGNSSATYFLLIGLPGLEAOFWLAFLPCLSYLAVLGNLTIIYVREHSIHHP 60  
DB 1 WVDNNGNSSATYFLLIGLPGLEAOFWLAFLPCLSYLAVLGNLTIIYVREHSIHHP 60  
QY MYIFCLMSGIDILISTSSMPKMLAIFWNSSTTIOFDACLCIQIPAHLSGMSSTVLLAM 120  
DB MYIFCLMSGIDILISTSSMPKMLAIFWNSSTTIOFDACLCIQIPAHLSGMSSTVLLAM 120  
QY 121 AFDRYVAICHPLRHATVTLTPRVTKIGVAAVRGALMAPLPVFIKQLPFCRSNITLSHSY 180  
DB 121 AFDRYVAICHPLRHATVTLTPRVTKIGVAAVRGALMAPLPVFIKQLPFCRSNITLSHSY 180  
QY 181 CHQDVMLACDDIRVNVVYGLIVISAIGDSLISFSTYLLIKTVLGLTREAOAKAFG 240  
DB 181 CHQDVMLACDDIRVNVVYGLIVISAIGDSLISFSTYLLIKTVLGLTREAOAKAFG 240  
QY 241 TCVSHVCAVFIFYVPFISLWVHRSKRDSPLPVILANIYLVLPVLPNIYGVTKKEI 300  
DB 241 TCVSHVCAVFIFYVPFISLWVHRSKRDSPLPVILANIYLVLPVLPNIYGVTKKEI 300  
QY 301 RORILRFHVATHASEP 317  
DB 301 RORILRFHVATHASEP 317

RESULT 2  
US-10-001-469-2880Sequence 2880, Application US/10001469  
Publication No. US20030091562A1

GENERAL INFORMATION:  
APPLICANT: JAKOBOVITS, AVA  
APPLICANT: RAYTANO, ARTHUR  
APPLICANT: AFAR, DANIEL  
APPLICANT: SAFERAN, DOUGLAS  
APPLICANT: HUBERT, RENE  
APPLICANT: PARIS, MARY  
APPLICANT: CHALLITA-EID, PTA  
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED  
FILE REFERENCE: 51158-20024.20  
CURRENT APPLICATION NUMBER: US/10/001,469  
PRIOR FILING DATE: 2002-05-20  
PRIOR APPLICATION NUMBER: 60/157,902  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: 60/291,118  
PRIOR FILING DATE: 2001-05-15  
PRIOR APPLICATION NUMBER: 09/680,728  
PRIOR FILING DATE: 2000-10-05  
NUMBER OF SEQ ID NOS: 2888  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2880  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: 10133A11 encoded amino  
US-10-001-469-2880

Query Match 100.0%; Score 1617; DB 15; Length 318;

Best Local Similarity 100.0%; Pred. No. 5,9e-150; Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVDNNGNSSATYFLLIGLPGLEAOFWLAFLPCLSYLAVLGNLTIIYVREHSIHHP 60  
DB 2 WVDNNGNSSATYFLLIGLPGLEAOFWLAFLPCLSYLAVLGNLTIIYVREHSIHHP 61

QY 61 MYIFCLMSGIDILISTSSMPKMLAIFWNSSTTIOFDACLCIQIPAHLSGMSSTVLLAM 120  
DB 62 MYIFCLMSGIDILISTSSMPKMLAIFWNSSTTIOFDACLCIQIPAHLSGMSSTVLLAM 121  
QY 121 AFDRYVAICHPLRHATVTLTPRVTKIGVAAVRGALMAPLPVFIKQLPFCRSNITLSHSY 180  
DB 122 AFDRYVAICHPLRHATVTLTPRVTKIGVAAVRGALMAPLPVFIKQLPFCRSNITLSHSY 181  
QY 181 CHQDVMLACDDIRVNVVYGLIVISAIGDSLISFSTYLLIKTVLGLTREAOAKAFG 240  
DB 182 CHQDVMLACDDIRVNVVYGLIVISAIGDSLISFSTYLLIKTVLGLTREAOAKAFG 241  
QY 241 TCVSHVCAVFIFYVPFISLWVHRSKRDSPLPVILANIYLVLPVLPNIYGVTKKEI 300  
DB 242 TCVSHVCAVFIFYVPFISLWVHRSKRDSPLPVILANIYLVLPVLPNIYGVTKKEI 301  
QY 301 RORILRFHVATHASEP 317  
DB 302 RORILRFHVATHASEP 318

RESULT 3  
US-10-295-027-873Sequence 873, Application US/10295027  
Publication No. US2003023250A1

GENERAL INFORMATION:  
APPLICANT: Afar, Daniel  
APPLICANT: Aziz, Natasha  
APPLICANT: Ginsberg, Wendy M.  
APPLICANT: Gish, Kurt C.  
APPLICANT: Glynn, Richard  
APPLICANT: Hevezl, Peter A.  
APPLICANT: Mack, David H.  
APPLICANT: Murray, Richard  
APPLICANT: Watson, Susan R.  
APPLICANT: Eos Biotechnology, Inc.  
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
FILE REFERENCE: 018501-012500US  
CURRENT APPLICATION NUMBER: US/10/295,027  
PRIOR FILING DATE: 2002-11-13  
PRIOR APPLICATION NUMBER: US 09/663,733  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: US 60/350,666  
PRIOR FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 60/335,394  
PRIOR FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: US 60/332,464  
PRIOR FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: US 60/334,393  
PRIOR FILING DATE: 2001-11-29  
PRIOR APPLICATION NUMBER: US 60/340,376  
PRIOR FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: US 60/347,211  
PRIOR FILING DATE: 2002-01-08  
PRIOR APPLICATION NUMBER: US 60/347,349  
PRIOR FILING DATE: 2002-01-10  
PRIOR APPLICATION NUMBER: US 60/355,250  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: US 60/356,714  
PRIOR FILING DATE: 2002-02-13  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1386  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 873  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens

US-10-295-027-873

Query Match 99.8%; Score 1614; DB 12; Length 317;  
Best Local Similarity 99.7%; Pred. No. 1.2e-149; Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPNGNESSATYFIIIGLPGLEBAQFWLAFPLCSYLAVGNLTIIYVREHSLHEP 60  
DB 1 MDPNGNESSATYFIIIGLPGLEBAQFWLAFPLCSYLAVGNLTIIYVREHSLHEP 60  
QY 61 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIFAIHSLSGMESTVLAM 120  
DB 61 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIFAIHSLSGMESTVLAM 120  
QY 121 AFRYVAICHLRHAATVLTLPRTKIGVAAYVGAALMAPLPVFIKQLPFCRSNIIISHSY 180  
DB 121 AFRYVAICHLRHAATVLTLPRTKIGVAAYVGAALMAPLPVFIKQLPFCRSNIIISHSY 180  
QY 181 CHQDVMKACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240  
DB 181 CHQDVMKACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240  
QY 241 TCVSHVCAVFIYVPIFGLSMWHRSKRDSPLPVILANIYLLVPPVLPVIYGVYKTKEI 300  
DB 241 TCVSHVCAVFIYVPIFGLSMWHRSKRDSPLPVILANIYLLVPPVLPVIYGVYKTKEI 300  
QY 301 RORILRLFHVATHASEP 317  
DB 301 RORILRLFHVATHASEP 317

## RESULT 4

US-10-292-798-746  
; Sequence 746, Application US/10292798  
; General Information:  
; GENERAL INFORMATION:  
; APPLICANT: SUMA, MAKIKO  
; APPLICANT: ASAI, KIYOSHI  
; APPLICANT: AKIYAMA, YUTAKA  
; APPLICANT: AUBRATANI, HIROYUKI  
; TITLE OF INVENTION: GUANOSINE TRI-PHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS  
; FILE REFERENCE: 084335/166  
; CURRENT APPLICATION NUMBER: US/10/292,798  
; PRIOR FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: 10/017,161  
; PRIOR FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: JP 2001-246789  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 2070  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 746  
; LENGTH: 317  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-292-798-746

Query Match 99.8%; Score 1614; DB 12; Length 317;  
Best Local Similarity 99.7%; Pred. No. 1,2e-149;  
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPNGNESSATYFIIIGLPGLEBAQFWLAFPLCSYLAVGNLTIIYVREHSLHEP 60  
DB 1 MDPNGNESSATYFIIIGLPGLEBAQFWLAFPLCSYLAVGNLTIIYVREHSLHEP 60  
QY 61 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIFAIHSLSGMESTVLAM 120  
DB 61 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIFAIHSLSGMESTVLAM 120  
QY 121 AFRYVAICHLRHAATVLTLPRTKIGVAAYVGAALMAPLPVFIKQLPFCRSNIIISHSY 180  
DB 121 AFRYVAICHLRHAATVLTLPRTKIGVAAYVGAALMAPLPVFIKQLPFCRSNIIISHSY 180  
QY 181 CHQDVMKACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240  
DB 181 CHQDVMKACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240  
QY 241 TCVSHVCAVFIYVPIFGLSMWHRSKRDSPLPVILANIYLLVPPVLPVIYGVYKTKEI 300  
DB 241 TCVSHVCAVFIYVPIFGLSMWHRSKRDSPLPVILANIYLLVPPVLPVIYGVYKTKEI 300

DB 241 TCVSHVCAVFIYVPIFGLSMWHRSKRDSPLPVILANIYLLVPPVLPVIYGVYKTKEI 300  
QY 301 RORILRLFHVATHASEP 317  
DB 301 RORILRLFHVATHASEP 317

## RESULT 5

US-09-759-143-920  
; Sequence 920, Application US/09759143  
; Patent No. US20020022248A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stoik, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darriek  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427023  
; CURRENT APPLICATION NUMBER: US/09/759,143  
; CURRENT FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 934  
; SOFTWARE: PaateSeq for Windows Version 3.0  
; SEQ ID NO 920  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-759-143-920

Query Match 99.8%; Score 1614; DB 9; Length 318;  
Best Local Similarity 99.7%; Pred. No. 1,2e-149;  
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPNGNESSATYFIIIGLPGLEBAQFWLAFPLCSYLAVGNLTIIYVREHSLHEP 60  
DB 2 MDPNGNESSATYFIIIGLPGLEBAQFWLAFPLCSYLAVGNLTIIYVREHSLHEP 61  
QY 61 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIFAIHSLSGMESTVLAM 120  
DB 62 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIFAIHSLSGMESTVLAM 121  
QY 121 AFRYVAICHLRHAATVLTLPRTKIGVAAYVGAALMAPLPVFIKQLPFCRSNIIISHSY 180  
DB 122 AFRYVAICHLRHAATVLTLPRTKIGVAAYVGAALMAPLPVFIKQLPFCRSNIIISHSY 181  
QY 181 CHQDVMKACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240  
DB 182 CHQDVMKACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 241  
QY 241 TCVSHVCAVFIYVPIFGLSMWHRSKRDSPLPVILANIYLLVPPVLPVIYGVYKTKEI 300  
DB 242 TCVSHVCAVFIYVPIFGLSMWHRSKRDSPLPVILANIYLLVPPVLPVIYGVYKTKEI 301  
QY 301 RORILRLFHVATHASEP 317  
DB 302 RORILRLFHVATHASEP 318

RESULT 6  
US-09-780-669-920

```
Sequence 920, Application us/09780669
Patent No. US20020051977A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Ranger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
APPLICANT: Hurai, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C24
CURRENT APPLICATION NUMBER: US/09/780,669
CURRENT FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 943
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 920
LENGTH: 318
TYPE: PRT
ORGANISM: Homo sapiens
US-09-780-669-920

Query Match          99.8%; Score 1614; DB 9; Length 318;
Best Local Similarity 99.7%; Pred. No. 1.2e-149;
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDPNNGESSATYFILLGLPGLBEOFWLAFPLCSLYLAVGNLTIIYVTRHSLHEP 60
DB 2 MVDPNNGESSATYFILLGLPGLBEOFWLAFPLCSLYLAVGNLTIIYVTRHSLHEP 61
QY MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQWFAHSLSGMESTVLLAM 120
DB 62 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQWFAHSLSGMESTVLLAM 121
QY 121 AFDRIYVAICHPRLRHATVLTLPRTVKIGVAAVVGALMAPLPVFIKQLPFCRSNLSHSY 180
DB 122 AFDRIYVAICHPRLRHATVLTLPRTVKIGVAAVVGALMAPLPVFIKQLPFCRSNLSHSY 181
QY 181 CHODVWKLACDDIRAVVYVGLIYIISAGIDSLISFVYLILKTVGLTREDAQKAFG 240
DB 182 CHODVWKLACDDIRAVVYVGLIYIISAGIDSLISFVYLILKTVGLTREDAQKAFG 241
QY 241 TCVSHVCAVFIYVFPFGLSMVHRFSKRDSPLPVILANTYLLVPVLPNIYGVKTKEI 300
DB 242 TCVSHVCAVFIYVFPFGLSMVHRFSKRDSPLPVILANTYLLVPVLPNIYGVKTKEI 301
QY 301 RORILRLFHVATHASEP 317
DB 302 RORILRLFHVATHASEP 318

RESULT 7
US-09-822-827-920
Sequence 920, Application US/09822827
Patent No. US20020081680A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
```

```
FILE REFERENCE: 210121.534C1
CURRENT APPLICATION NUMBER: US/09/822,827
CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 920
LENGTH: 318
TYPE: PRT
ORGANISM: Homo sapiens
US-09-822-827-920

Query Match          99.8%; Score 1614; DB 9; Length 318;
Best Local Similarity 99.7%; Pred. No. 1.2e-149;
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDPNNGESSATYFILLGLPGLBEOFWLAFPLCSLYLAVGNLTIIYVTRHSLHEP 60
DB 2 MVDPNNGESSATYFILLGLPGLBEOFWLAFPLCSLYLAVGNLTIIYVTRHSLHEP 61
QY MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQWFAHSLSGMESTVLLAM 120
DB 62 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQWFAHSLSGMESTVLLAM 121
QY 121 AFDRIYVAICHPRLRHATVLTLPRTVKIGVAAVVGALMAPLPVFIKQLPFCRSNLSHSY 180
DB 122 AFDRIYVAICHPRLRHATVLTLPRTVKIGVAAVVGALMAPLPVFIKQLPFCRSNLSHSY 181
QY 181 CHODVWKLACDDIRAVVYVGLIYIISAGIDSLISFVYLILKTVGLTREDAQKAFG 240
DB 182 CHODVWKLACDDIRAVVYVGLIYIISAGIDSLISFVYLILKTVGLTREDAQKAFG 241
QY 241 TCVSHVCAVFIYVFPFGLSMVHRFSKRDSPLPVILANTYLLVPVLPNIYGVKTKEI 300
DB 242 TCVSHVCAVFIYVFPFGLSMVHRFSKRDSPLPVILANTYLLVPVLPNIYGVKTKEI 301
QY 301 RORILRLFHVATHASEP 317
DB 302 RORILRLFHVATHASEP 318

RESULT 8
US-09-886-055-91
Sequence 91, Application US/09886055
Patent No. US20020132273A1
GENERAL INFORMATION:
APPLICANT: STREYER, LOBERT
APPLICANT: ZOZULYA, SERGEY
TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
FILE REFERENCE: 078003-0277150
CURRENT APPLICATION NUMBER: US/09/886,055
CURRENT FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: 60/213,812
PRIOR FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 522
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 91
LENGTH: 318
TYPE: PRT
ORGANISM: Homo sapiens
US-09-886-055-91

Query Match          99.8%; Score 1614; DB 10; Length 318;
Best Local Similarity 99.7%; Pred. No. 1.2e-149;
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDPNNGESSATYFILLGLPGLBEOFWLAFPLCSLYLAVGNLTIIYVTRHSLHEP 60
DB 2 MVDPNNGESSATYFILLGLPGLBEOFWLAFPLCSLYLAVGNLTIIYVTRHSLHEP 61
QY MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQWFAHSLSGMESTVLLAM 120
DB 62 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQWFAHSLSGMESTVLLAM 121
```

```
QY 121 AFRRYVAICHPRLHATVLTLPRTVKIGVAAVVRGAALMAELPVFIKQLPFCRSNLSHSY 180
DB 122 AFRRYVAICHPRLHATVLTLPRTVKIGVAAVVRGAALMAELPVFIKQLPFCRSNLSHSY 181
QY 181 CLHODVWKLACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240
DB 182 CLHODVWKLACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 241
QY 241 TCVSHVCAVFIFFYPFICLSMWRHFSKRSDPLPVILANILYLVPLNPIVYGVTKEI 300
DB 242 TCVSHVCAVFIFFYPFICLSMWRHFSKRSDPLPVILANILYLVPLNPIVYGVTKEI 301
QY 301 RORILRFLHVATHASEP 317
DB 302 RORILRFLHVATHASEP 318
```

## RESULT 9

```
US-09-895-793-920
; Sequence 920, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 920
; LENGTH: 318
; TYPE: PRS
; ORGANISM: Homo sapiens
US-09-895-793-920
```

```
Query Match 99.8%; Score 1614; DB 10; Length 318;
Best Local Similarity 99.7%; Pred. No. 1.2e-149;
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MDPNGNESATYFLLIGLPGLEAOFWLAFLCSLYLAVGNLTITIVTRTHSHLP 60
DB 2 MDPNGNESATYFLLIGLPGLEAOFWLAFLCSLYLAVGNLTITIVTRTHSHLP 61
QY 61 MYIFCLMSGIDILISTSSMPKMLAIFWNSSTTIQFDACLLQMFALHSLSGMSSTVLLAM 120
DB 62 MYIFCLMSGIDILISTSSMPKMLAIFWNSSTTIQFDACLLQMFALHSLSGMSSTVLLAM 121
QY 121 AFRRYVAICHPRLHATVLTLPRTVKIGVAAVVRGAALMAELPVFIKQLPFCRSNLSHSY 180
DB 122 AFRRYVAICHPRLHATVLTLPRTVKIGVAAVVRGAALMAELPVFIKQLPFCRSNLSHSY 181
```

```
QY 181 CLHODVWKLACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240
DB 182 CLHODVWKLACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 241
QY 241 TCVSHVCAVFIFFYPFICLSMWRHFSKRSDPLPVILANILYLVPLNPIVYGVTKEI 300
DB 242 TCVSHVCAVFIFFYPFICLSMWRHFSKRSDPLPVILANILYLVPLNPIVYGVTKEI 301
QY 301 RORILRFLHVATHASEP 317
DB 302 RORILRFLHVATHASEP 318
```

## RESULT 10

```
US-09-895-814-920
; Sequence 920, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 920
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 920
; LENGTH: 318
; TYPE: PRS
; ORGANISM: Homo sapiens
US-09-895-814-920
```

```
Query Match 99.8%; Score 1614; DB 10; Length 318;
Best Local Similarity 99.7%; Pred. No. 1.2e-149;
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MDPNGNESATYFLLIGLPGLEAOFWLAFLCSLYLAVGNLTITIVTRTHSHLP 60
DB 2 MDPNGNESATYFLLIGLPGLEAOFWLAFLCSLYLAVGNLTITIVTRTHSHLP 61
QY 61 MYIFCLMSGIDILISTSSMPKMLAIFWNSSTTIQFDACLLQMFALHSLSGMSSTVLLAM 120
DB 62 MYIFCLMSGIDILISTSSMPKMLAIFWNSSTTIQFDACLLQMFALHSLSGMSSTVLLAM 121
QY 121 AFRRYVAICHPRLHATVLTLPRTVKIGVAAVVRGAALMAELPVFIKQLPFCRSNLSHSY 180
DB 122 AFRRYVAICHPRLHATVLTLPRTVKIGVAAVVRGAALMAELPVFIKQLPFCRSNLSHSY 181
QY 181 CLHODVWKLACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240
DB 182 CLHODVWKLACDDIRVNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 241
```

Tue Feb 10 06:04:27 2004

us-10-017-066a-2. rapb

Page 6

QY 241 TCVSHVCAVFIFFVYFGLSMWHRFSKRDSPLPVIIANTYLLVPPVLANPIYGVKTKKEI 300  
DB 242 TCVSHVCAVFIFFVYFGLSMWHRFSKRDSPLPVIIANTYLLVPPVLANPIYGVKTKKEI 301  
QY 301 RORILRLFHVATHASEP 317  
DB 302 RORILRLFHVATHASEP 318

RESULT 11  
US-09-966-459A-2  
; Sequence 2, Application US/09966459A  
; Publication No. US20030022237A1  
; GENERAL INFORMATION:  
; APPLICANT: FEDER, J.N.  
; APPLICANT: MINTIER, G.  
; APPLICANT: RAMANATHAN, C.S.  
; APPLICANT: HAWKEN, D.R.  
; APPLICANT: CACCACE, A.  
; APPLICANT: BARBER, L.  
; APPLICANT: KORNAK, M.G.  
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGRPMY4,  
; FILE REFERENCE: D0039NP  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: US/09/966,459A  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: 60/235,833  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/261,776  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/305,351  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/313,202  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-966-459A-2

Query Match 99.8%; Score 1614; DB 11; Length 318;  
Best Local Similarity 99.7%; Pred. No. 1.2e-149;  
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYIFCLMSGIDILISTSSMPKMLAIFWENSTTIQFDACLLQIFAHISGMESTVLLAM 120  
DB 2 MYIFCLMSGIDILISTSSMPKMLAIFWENSTTIQFDACLLQIFAHISGMESTVLLAM 121  
QY 61 MYIFCLMSGIDILISTSSMPKMLAIFWENSTTIQFDACLLQIFAHISGMESTVLLAM 120  
DB 62 MYIFCLMSGIDILISTSSMPKMLAIFWENSTTIQFDACLLQIFAHISGMESTVLLAM 121  
QY 121 AFDRYVAICPLRHATVLTLPRTYKIGVAAVVRGAALMAPLPVFIKQLPFCRSNIIISHSY 180  
DB 122 AFDRYVAICPLRHATVLTLPRTYKIGVAAVVRGAALMAPLPVFIKQLPFCRSNIIISHSY 181  
QY 181 CLHODVWKLACDDIRVNVVYGLIYISAGLDSILISFSYLLIKTVLGLTREQAQAFG 240  
DB 182 CLHODVWKLACDDIRVNVVYGLIYISAGLDSILISFSYLLIKTVLGLTREQAQAFG 241  
QY 241 TCVSHVCAVFIFFVYFGLSMWHRFSKRDSPLPVIIANTYLLVPPVLANPIYGVKTKKEI 300  
DB 242 TCVSHVCAVFIFFVYFGLSMWHRFSKRDSPLPVIIANTYLLVPPVLANPIYGVKTKKEI 301  
QY 301 RORILRLFHVATHASEP 317  
DB 302 RORILRLFHVATHASEP 318

RESULT 12

US-09-804-291-91  
; Sequence 91, Application US/09804291  
; Publication No. US2003008059A1  
; GENERAL INFORMATION:  
; APPLICANT: ZOZULA, SERGEY  
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME  
; FILE REFERENCE: P 0278005  
; CURRENT FILING DATE: 2001-03-13  
; PRIOR APPLICATION NUMBER: US/09/804,291  
; PRIOR FILING DATE: 2000-03-13  
; PRIOR APPLICATION NUMBER: 60/188,914  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,033  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/198,474  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/199,335  
; PRIOR FILING DATE: 2000-04-24  
; PRIOR APPLICATION NUMBER: 60/207,702  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/213,849  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/226,534  
; PRIOR FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: 60/230,732  
; PRIOR FILING DATE: 2000-09-07  
; PRIOR APPLICATION NUMBER: 60/266,862  
; PRIOR FILING DATE: 2001-02-07  
; NUMBER OF SEQ ID NOS: 529  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 91  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-804-291-91

Query Match 99.8%; Score 1614; DB 11; Length 318;  
Best Local Similarity 99.7%; Pred. No. 1.2e-149;  
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYIFCLMSGIDILISTSSMPKMLAIFWENSTTIQFDACLLQIFAHISGMESTVLLAM 120  
DB 2 MYIFCLMSGIDILISTSSMPKMLAIFWENSTTIQFDACLLQIFAHISGMESTVLLAM 121  
QY 61 MYIFCLMSGIDILISTSSMPKMLAIFWENSTTIQFDACLLQIFAHISGMESTVLLAM 120  
DB 62 MYIFCLMSGIDILISTSSMPKMLAIFWENSTTIQFDACLLQIFAHISGMESTVLLAM 121  
QY 121 AFDRYVAICPLRHATVLTLPRTYKIGVAAVVRGAALMAPLPVFIKQLPFCRSNIIISHSY 180  
DB 122 AFDRYVAICPLRHATVLTLPRTYKIGVAAVVRGAALMAPLPVFIKQLPFCRSNIIISHSY 181  
QY 181 CLHODVWKLACDDIRVNVVYGLIYISAGLDSILISFSYLLIKTVLGLTREQAQAFG 240  
DB 182 CLHODVWKLACDDIRVNVVYGLIYISAGLDSILISFSYLLIKTVLGLTREQAQAFG 241  
QY 241 TCVSHVCAVFIFFVYFGLSMWHRFSKRDSPLPVIIANTYLLVPPVLANPIYGVKTKKEI 300  
DB 242 TCVSHVCAVFIFFVYFGLSMWHRFSKRDSPLPVIIANTYLLVPPVLANPIYGVKTKKEI 301  
QY 301 RORILRLFHVATHASEP 317  
DB 302 RORILRLFHVATHASEP 318

RESULT 13  
US-10-017-161-868  
; Sequence 868, Application US/10017161  
; Publication No. US2003014366A1  
; GENERAL INFORMATION:  
; APPLICANT: SUWA, MAKIKO  
; APPLICANT: ASAI, KIYOSHI  
; APPLICANT: AKIYAMA, YUTAKA  
; APPLICANT: ABURATANI, HIROYUKI

TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS  
FILE REFERENCE: 084335/0152  
CURRENT APPLICATION NUMBER: US/10/017,161  
CURRENT FILING DATE: 2002-12-18  
PRIOR APPLICATION NUMBER: JP 2001/246789  
PRIOR FILING DATE: 2001-06-18  
NUMBER OF SEQ ID NOS: 2430  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 868  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-017-161-868

Query Match 99.8%; Score 1614; DB 12; Length 318;  
Best Local Similarity 99.7%; Pred. No. 1.2e-149;  
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPNGESSATYFIILIGLGEAOFWLAFLCSYLAVGNLTIIYVREHSLHEP 60  
DB 2 MDPNGESSATYFIILIGLGEAOFWLAFLCSYLAVGNLTIIYVREHSLHEP 61  
QY 61 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIIFAHSLSGMESTVLLAM 120  
DB 62 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIIFAHSLSGMESTVLLAM 121  
QY 121 AFDRIYALCHPLRHATVLTLPRTKIGVAAVVGALMLAPLPIFIKQLPFCRSNLSHSY 180  
DB 122 AFDRIYALCHPLRHATVLTLPRTKIGVAAVVGALMLAPLPIFIKQLPFCRSNLSHSY 181  
QY 181 CLHQDVAKLACDDIRNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240  
DB 182 CLHQDVAKLACDDIRNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 241  
QY 241 TCVSHVCAVFIYVPFIGLSMVHRSKRDSPLPVILANIYLLVPVLPVYGVTKKEI 300  
DB 242 TCVSHVCAVFIYVPFIGLSMVHRSKRDSPLPVILANIYLLVPVLPVYGVTKKEI 301  
QY 301 RORILRLFHVATHASEP 317  
DB 302 RORILRLFHVATHASEP 318

RESULT 14  
US-10-144-678A-920  
Sequence 920, Application US/10144678A  
Publication No. US20030157089A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, David C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yugui  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Ranger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Derrick  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skelley, Yasir A. W.  
APPLICANT: Hepler, William T.  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals Y de Basols, Carlotia  
APPLICANT: Roy, Teresa M.  
APPLICANT: Watanabe, Yoshihiro  
APPLICANT: Deng, Ta

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C28  
CURRENT APPLICATION NUMBER: US/10/144,678A  
CURRENT FILING DATE: 2002-08-12  
NUMBER OF SEQ ID NOS: 1033  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 920  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-144-678A-920

Query Match 99.8%; Score 1614; DB 12; Length 318;  
Best Local Similarity 99.7%; Pred. No. 1.2e-149;  
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPNGESSATYFIILIGLGEAOFWLAFLCSYLAVGNLTIIYVREHSLHEP 60  
DB 2 MDPNGESSATYFIILIGLGEAOFWLAFLCSYLAVGNLTIIYVREHSLHEP 61  
QY 61 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIIFAHSLSGMESTVLLAM 120  
DB 62 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIIFAHSLSGMESTVLLAM 121  
QY 121 AFDRIYALCHPLRHATVLTLPRTKIGVAAVVGALMLAPLPIFIKQLPFCRSNLSHSY 180  
DB 122 AFDRIYALCHPLRHATVLTLPRTKIGVAAVVGALMLAPLPIFIKQLPFCRSNLSHSY 181  
QY 181 CLHQDVAKLACDDIRNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 240  
DB 182 CLHQDVAKLACDDIRNVVYGLIIVISAIGDLSLISFSYLLIKTVLGLTREAOAKAFG 241  
QY 241 TCVSHVCAVFIYVPFIGLSMVHRSKRDSPLPVILANIYLLVPVLPVYGVTKKEI 300  
DB 242 TCVSHVCAVFIYVPFIGLSMVHRSKRDSPLPVILANIYLLVPVLPVYGVTKKEI 301  
QY 301 RORILRLFHVATHASEP 317  
DB 302 RORILRLFHVATHASEP 318

RESULT 15  
US-10-294-025-920  
Sequence 920, Application US/10294025  
Publication No. US20030185830A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Stolk, John A.  
APPLICANT: Kalos, Michael D.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C29  
CURRENT APPLICATION NUMBER: US/10/294,025  
CURRENT FILING DATE: 2002-11-12  
NUMBER OF SEQ ID NOS: 1038  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 920  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-294-025-920

Query Match 99.8%; Score 1614; DB 12; Length 318;  
Best Local Similarity 99.7%; Pred. No. 1.2e-149;  
Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPNGESSATYFIILIGLGEAOFWLAFLCSYLAVGNLTIIYVREHSLHEP 60  
DB 2 MDPNGESSATYFIILIGLGEAOFWLAFLCSYLAVGNLTIIYVREHSLHEP 61  
QY 61 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIIFAHSLSGMESTVLLAM 120  
DB 62 MYFLCMLSGIDILISTSSMPKMLAFWENSTTIQDACLQIIFAHSLSGMESTVLLAM 121

Qy	121	AEDRYVAICHPLRHATVLTLPRTVKIGVAAVVRGAMAPLPVFIKQLPFCRSNIISHSY	180
Db	122	AFDRYVAICHPLRHATVLTLPRTVKIGVAAVVRGAMAPLPVFIKQLPFCRSNIISHSY	181
Qy	181	CIHODVMKLAACDDIRVNVVYGLIVISAGIDSLISFSYLLIKTVLGITREAOAKFG	240
Db	182	CIHODVMKLAACDDIRVNVVYGLIVISAGIDSLISFSYLLIKTVLGITREAOAKFG	241
Qy	241	TCVSHVCAVFIYVPFGLSMVHRFSKRDSPLPVILANIYLLVPPYLPNIYGVYKXKEI	300
Db	242	TCVSHVCAVFIYVPFGLSMVHRFSKRDSPLPVILANIYLLVPPYLPNIYGVYKXKEI	301
Qy	301	RQRLRLFHVATHASEP	317
Db	302	RQRLRLFHVATHASEP	318

Search completed: February 9, 2004, 16:30:47  
 Job time : 35 secs



Result No.	Score	Query Match	Length	DB	ID	Description
1	1614	99.8	317	4	Q8TCH6	Q8TCH6 homo sapien
2	1512	93.5	320	11	Q8VQZ7	Q8VQZ7 mus musculus
3	985	60.9	320	11	Q8VBV5	Q8VBV5 mus musculus
4	965.5	59.7	322	11	Q8VFP9	Q8VFP9 mus musculus
5	937.5	58.0	324	4	Q8NGF3	Q8NGF3 homo sapien
6	914.5	56.6	315	11	Q8VYH5	Q8VYH5 mus musculus
7	910.5	55.7	312	11	Q8VH11	Q8VH11 mus musculus
8	892.5	55.2	314	4	Q8NGK1	Q8NGK1 homo sapien
9	877	54.2	313	11	Q8VH17	Q8VH17 mus musculus
10	875.5	54.1	314	11	Q8VQZ6	Q8VQZ6 mus musculus
11	873.5	54.0	314	11	Q8VQZ2	Q8VQZ2 mus musculus
12	873.5	54.0	318	11	Q8VQZ3	Q8VQZ3 mus musculus
13	867.5	53.3	317	11	Q8VH01	Q8VH01 mus musculus
14	857.5	53.0	319	11	Q8VGM0	Q8VGM0 mus musculus
15	857.5	53.0	319	11	Q8VQZ8	Q8VQZ8 mus musculus
16	854	52.8	317	11	Q8VH04	Q8VH04 mus musculus

17	852.5	52.7	319	11	Q8EC07	Q8EC07	mus	musculu
18	852	52.7	317	11	Q8VEX9	Q8VEX9	mus	musculu
19	844.5	52.2	319	11	Q8VY9	Q8VY9	mus	musculu
20	844.5	52.0	329	4	Q8NH57	Q8NH57	homo	sapien
21	838.5	51.9	312	11	Q8VXK6	Q8VXK6	mus	musculu
22	835.5	51.7	314	11	Q8VH15	Q8VH15	mus	musculu
23	835	51.6	330	11	Q8VFO7	Q8VFO7	mus	musculu
24	834	51.6	316	11	Q8VX8	Q8VX8	mus	musculu
25	833.5	51.5	317	11	Q8VX8	Q8VX8	mus	musculu
26	831.5	51.4	319	11	Q8VY6	Q8VY6	mus	musculu
27	831.5	51.4	327	11	Q8VFX28	Q8VFX28	mus	musculu
28	831	51.4	312	11	Q8VH12	Q8VH12	mus	musculu
29	829	51.3	317	11	Q8VQV9	Q8VQV9	mus	musculu
30	829	51.3	318	11	Q8VH16	Q8VH16	mus	musculu
31	827.5	51.2	315	11	Q8VGB5	Q8VGB5	mus	musculu
32	827.5	51.2	315	4	Q8NG35	Q8NG35	homo	sapien
33	826	51.1	316	11	Q8VQ79	Q8VQ79	mus	musculu
34	824	51.0	317	4	Q8NGK5	Q8NGK5	homo	sapien
35	824	51.0	321	4	Q8NGK1	Q8NGK1	homo	sapien
36	823.5	50.9	329	11	Q8VX5	Q8VX5	mus	musculu
37	823	50.9	325	11	Q8VYX9	Q8VYX9	mus	musculu
38	822.5	50.9	316	11	Q8VQZ0	Q8VQZ0	mus	musculu
39	818.5	50.6	312	11	Q8VFO2	Q8VFO2	mus	musculu
40	817	50.5	313	4	Q8NGJ6	Q8NGJ6	homo	sapien
41	816	50.5	319	13	Q8VH55	Q8VH55	gallus	gall
42	814.5	50.4	313	11	Q8VY4	Q8VY4	mus	musculu
43	814	50.3	315	11	Q8VQZ3	Q8VQZ3	mus	musculu
44	812.5	50.2	314	4	Q8NGK3	Q8NGK3	homo	sapien
45	811.5	50.2	314	11	Q8VH21	Q8VH21	mus	musculu

## ALIGNMENTS

## RESULT 1

ID	Q8TCB6	PRELIMINARY;	PRT;	317 AA.
----	--------	--------------	------	---------

DT 01-JUN-2002 (Tremblayrel. 21, Created)  
DT 01-JUN-2002 (Tremblayrel. 21, Last sequence update)  
DT 01-MAR-2003 (Tremblayrel. 23, Last annotation update)  
D3 Hypochemical protein (Seven transmembrane helix receptor).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Lung;  
RC Strausberg R.;  
RU Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.  
RA Suwa M., Sato T., Okouchi I., Arima M., Futami K., Matsumoto S.,  
RA Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;  
RT "genome-wide discovery and analysis of human seven transmembrane helix  
RT receptor genes."; <http://www.ncbi.nlm.nih.gov/pubmed/11500000>.  
RU Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR000276; GPCR\_Rhodopsin.  
DR Pfam: P00000.7cm1.1.  
DR PROSITE: PS00337; G\_PROTEIN\_RECP\_F1\_1.1.  
DR PROSITE: PS0062; G\_PROTEIN\_RECP\_F1\_2.1.  
KW Hypothetical protein; Receptor; Transmembrane.  
SQ SEQUENCE 317 AA; 35271 MW; 8C7293AA7FBCA35C CRC64

Query Match	99.8%	Score 1614	DB 4	Length 317
Best Local Similarity	99.7%	Pred. No. 1.6e-135		
Matches 316; Conservative	1	Mismatches 0	Indels 0	Gaps 0

QY I MVDPNGNESSATYFILIGLPGLEEAFWLAFLPGLSYLLIYAVLGNLTIIYIVRTESLSHEP 600

```

Db      1  MVDNGNSSATYITLLGLPGLBEAQFWLAFPLCSLYLIVLGNLTITIVYVREHSLHEP 60
Qy      61  MYIFLCMLSGIDILISTSSMPKMLAFWENSTTIQPDACILQIFAIHSLSGMESTVLLAM 120
Db      61  MYIFLCMLSGIDILISTSSMPKMLAFWENSTTIQPDACILQIFAIHSLSGMESTVLLAM 120
Qy      121  AFDRYVAICHPRLRAVTLTPRVTIKIGAAVVRGAALMAPLPFIRKLPFCRSNIIISHSY 180
Db      121  AFDRYVAICHPRLRAVTLTPRVTIKIGAAVVRGAALMAPLPFIRKLPFCRSNIIISHSY 180
Qy      181  CLHODVWKLACDDIRVNVVYGLIYISALIGDLSILISFSYLLIKTVLGLTREAOQAKAG 240
Db      181  CLHODVWKLACDDIRVNVVYGLIYISALIGDLSILISFSYLLIKTVLGLTREAOQAKAG 240
Qy      241  TCVSHVCAVFIFFYVFFIGLSWVHRFSKRDSPLPYIANTIVLVPVLPNTIVYGVTKKEI 300
Db      241  TCVSHVCAVFIFFYVFFIGLSWVHRFSKRDSPLPYIANTIVLVPVLPNTIVYGVTKKEI 300
Qy      301  RORILRLFHVATTHASEP 317
Db      301  RORILRLFHVATTHASEP 317

RESULT 2
ID      08VGZ7  PRELIMINARY;  PRT;  317 AA.
AC      08VGZ7;
DT      01-MAR-2002 (TREMBlrel. 20, Created)
DT      01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      Olfactory receptor MOR18-1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_Taxid=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Zhang X., Firestein S.J.;
RT      "The olfactory receptor gene superfamily of the mouse.";
RL      Nat. Neurosci. 0:0-0(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Adams M.;
RT      Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL      EMBL; AY072993; AAL06056.1; -.
DR      InterPro; IPR000276; GPCR_Rhodopsn.
DR      Pfam; PF00001; 7tm_1; 1.
DR      PRINTS; PR00237; GPCR_RHODOPSIN.
DR      PROSITE; PS00237; G_PROTEIN_RECPT_F1_1; 1.
DR      PROSITE; PS50262; G_PROTEIN_RECPT_F1_2; 1.
KW      Receptor.
SQ      SEQUENCE 317 AA; 35463 MW; 5A0E59E229852534 CRC64;

Query Match 93.5%; Score 1512; DB 11; Length 317;
Best Local Similarity 93.4%; Pred. No. 1.9e-126;
Matches 295; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

```

```

Qy      241  TCVSHVCAVFIFFYVFFIGLSWVHRFSKRDSPLPYIANTIVLVPVLPNTIVYGVTKKEI 300
Db      241  TCVSHVCAVFIFFYVFFIGLSWVHRFSKRDSPLPYIANTIVLVPVLPNTIVYGVTKKEI 300
Qy      301  RORILRLFHVATTHASEP 316
Db      301  RORILRLFHVATTHASEP 316

RESULT 3
ID      08VEV9  PRELIMINARY;  PRT;  320 AA.
AC      08VEV9;
DT      01-MAR-2002 (TREMBlrel. 20, Created)
DT      01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      Olfactory receptor MOR18-2 (Prostate-specific G protein-coupled
DE      receptor RALc).
GN      OLF78.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_Taxid=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Zhang X., Firestein S.J.;
RT      "The olfactory receptor gene superfamily of the mouse.";
RL      Nat. Neurosci. 0:0-0(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Adams M.;
RT      Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL      [3]
RP      SEQUENCE FROM N.A.
RA      STRAIN-BALB/c;
RX      MEDLINE=21564169; PubMed=11707321;
RA      Yuan T.T., Toy P., McClary J.A., Lin R.J., Miyamoto N.G.,
RA      Kretschmer P.J.;
RT      "Cloning and genetic characterization of an evolutionarily conserved
RT      human olfactory receptor that is differentially expressed across
RT      species.";
RL      Gene 278:41-51(2001).
RN      [4]
RP      SEQUENCE FROM N.A.
RA      STRAIN=C57BL/6J; TISSUE=Cerebellum, and Skin;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium,
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573(2002).
DR      EMBL; AY073011; AAL0674.1; -.
DR      EMBL; AF378854; AAL35109.1; -.
DR      EMBL; AK028467; BAC25966.1; -.
DR      EMBL; AK036356; BAC29396.1; -.
DR      MGD; MGI:2157548; Olf78.
DR      InterPro; IPR000276; GPCR_Rhodopsn.
DR      Pfam; PF00001; 7tm_1; 1.
DR      PRINTS; PR00237; GPCR_RHODOPSIN.
DR      PROSITE; PS00237; G_PROTEIN_RECPT_F1_1; 1.
DR      PROSITE; PS50262; G_PROTEIN_RECPT_F1_2; 1.
KW      Receptor.
SQ      SEQUENCE 320 AA; 35577 MW; 7161ACAF4328959 CRC64;

Query Match 60.9%; Score 985; DB 11; Length 320;
Best Local Similarity 59.8%; Pred. No. 1.3e-79;
Matches 183; Conservative 53; Mismatches 68; Indels 2; Gaps 2;

```

```
Db 64 MAAIDIALSTSTMPKILAFWDSREITFDACIAOMPFHILSLAESTILLAMAFDRYV 123
Qy 127 AICHPLRHATVTLPLPRYTKIGVAAYVRGALMALPLVPFIKOLPECRSNILSHSYCLHODV 186
Db 124 AICHPLRHAAVNLNTYVQIGMVALVRGSLFFPPLILIKRLAFCHSNVLSHSYCVHODV 183
Qy 187 MKIACDDIRVNVVYGLIVIIISALGDSLISFSYLLILKTVLGL-TREAQAKAFGTCVSH 245
Db 184 MKIAYDTLPNVVYGLTALILVMGVMEFISLSYFLIIRTVQLPEKSRRAKAFGTCVSH 243
Qy 246 VCAVFIFYPFIFGLSMVHRFSKRRDSPPLVILANITLVLPVLANIPIYGVTKREIRORTL 305
Db 244 ISVVAIFVYVPLIGLSVHRFGNSLDPTVHVMGDVYLLPVPINPIYGAKTQIRTRVL 303
Qy 306 RLPHVA 311
Db 304 AMFKIS 309

RESULT 4
Q8VF09 PRELIMINARY; PRT; 322 AA.
AC Q8VF09;
DT 01-MAR-2002 (T-REMBLrel. 20, Created)
DT 01-MAR-2002 (T-REMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (T-REMBLrel. 23, Last annotation update)
DE Olfactory receptor MOR18-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang X., Firestein S.J.;
RT "The olfactory receptor gene superfamily of the mouse.";
RL Nat. Neurosci. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Adams M.;
RT Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
RL EMBL: AY073732; AAL61395.1;
DR InterPro: IPR00276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHOPOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 322 AA; 35612 MW; D40302782D4B2PFE CRC64;

Query Match 59.7%; Score 965.5; DB 11; Length 322;
Best Local Similarity 59.0%; Pred. No. 7.3e-78;
Matches 181; Conservative 57; Mismatches 66; Indels 3; Gaps 3;
```

```
Qy 302 QRILRF 308
Db 311 SRVIRMF 317

RESULT 5
Q8NGF3 PRELIMINARY; PRT; 324 AA.
AC Q8NGF3;
DT 01-OCT-2002 (T-REMBLrel. 22, Created)
DT 01-OCT-2002 (T-REMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-REMBLrel. 23, Last annotation update)
DE Seven transmembrane helix receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT "Genome-wide discovery and analysis of human seven transmembrane helix
RT receptor genes.";
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB065855; BAC06073.1;
DR InterPro: IPR00276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHOPOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor; Transmembrane.
SQ SEQUENCE 324 AA; 35839 MW; 52401F86565E3BD1 CRC64;

Query Match 58.0%; Score 937.5; DB 4; Length 324;
Best Local Similarity 58.5%; Pred. No. 2.3e-75;
Matches 179; Conservative 52; Mismatches 72; Indels 3; Gaps 3;
```

```
Qy 5 NNGESSATYFILIGLRL-EEAOFMLAFPLCSLYLAVLGNLTITIVYREHSHLEPMYI 63
Db 15 NGMLVHAAYPLLVGIPGLPTTFHMLAFPLCFMYAATLGNLTIVILIRERLHEPMYL 74
Qy 64 FLCLMSGIDILISTSSNPKMLAFWNSSTTIQPDACILQIFAIHSGMESVYLLAMAFD 123
Db 75 FLMLSTIDVLSSIMPKWASFLMGIGIEIBENICLQMLFHLHSAVSAYLLAMAFD 134
Qy 124 RYVAICHPLRHATVTLPLPRYTKIGVAAYVRGALMALPLVPFIKOLPECRSNILSHSYCLH 183
Db 135 RYVAICHPLRHASVLTGCTVAKIGLSALTRGFVFPFPLIKMLSYCQTHVTHSFCLH 194
Qy 184 QDVKIACDDIRVNVVYGLIVIIISALGDSLISFSYLLILKTVLGL-TREAQAKAFGTC 242
Db 195 QDLMKLSCTDTRNVVYGLFIIISVGVDSLFIFGFIYIILNAVLELSRRALTKAFNFC 254
Qy 243 VSHVCAVFIFYPFIFGLSMVHRFSKRRDSPPLVILANITLVLPVLANIPIYGVTKREIRQ 302
Db 255 ISHLCAVLAFFVYVPLIGLSVHRFG-GLTSLHLYVMANTYLLPVPVNPVYGAKTKEICS 313
Qy 303 RLRLRF 308
Db 314 RVLQMF 319

RESULT 6
Q8VH05 PRELIMINARY; PRT; 315 AA.
AC Q8VH05;
DT 01-MAR-2002 (T-REMBLrel. 20, Created)
DT 01-MAR-2002 (T-REMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (T-REMBLrel. 23, Last annotation update)
DE Olfactory receptor MOR10-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



```

Db      68 LSMALIDLGSLCTLPVLGI FVWGAREISHDCAFLQFFIHCFSSLESSVLLISMAFDR 127
QY      125 YVALCHPLRHATVTLTPRVTKIGVAAYVGAALMAFLPEVFIKQLPFCRSNIISHSYCLHQ 184
Db      128 FVALCHPLRHATVTLTPRVTKIGVAAYVGAALMAFLPEVFIKQLPFCRSNIISHSYCLHQ 187
QY      185 DVMKLACDDIRVVNVYGLIIVISAIGDLSLISFYLILKTVTGLTREACA-KAFETCV 243
Db      188 EVMKLACADMRANSIYGMFIVISVIGDLSLISFYLILKTVTGLTREACA-KAFETCV 247
QY      244 SHVCAVFIFYVPFISLWVHRFSKRDSPLPVILANIYLVPLVPLNPIYGVKTKRIOR 303
Db      248 SHICAVLLFYTPMIGLSVHRFGKAPHLVQVWGMFVLLFPVPMNIIVISVKTQIKRDR 307
QY      304 ILRLP 308
Db      308 VTHAF 312

```

## RESULT 9

```

Q8VH17 PRELIMINARY; PRT; 313 AA.
ID      08VH17
AC      01-MAR-2002 (Tremblrel. 20, Created)
DT      01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT      01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE      Olfactory receptor MOR7-1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN      [1]
RP      NCBI_TaxId=10090;
RA      Zhang X., Firestein S.J.;
RT      "The olfactory receptor gene superfamily of the mouse.";
RL      Nat. Neurosci. 0:0-0(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Adams M.;
RL      Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR      EMBL; AY072971; AAL60634.1; -
DR      InterPro; IPR000276; GPCR_Rhodopsn.
DR      Pfam; PF00001; 7tm_1; 1.
DR      PRINTS; PR00237; GPCR_RHODOPSN.
DR      PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR      PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
KM      Receptor.
SQ      SEQUENCE 313 AA; 35682 MW; 39476E03FD201DEC CRC64;

```

Query Match 54.2%; Score 877; DB 11; Length 313;

Best Local Similarity 51.8%; Pred. No. 5.2e-70; Matches 158; Conservative 72; Mismatches 73; Indels 2; Gaps 2;

```

QY      5 NGENESATYFIIIGPGLBEAOFMLAPLCSYLYLAVGNLTIIYIVTEHSLHEPMYIF 64
Db      6 NSLSLQKATFF-LTGSGLEBERHGWISIFCSYLYLVIIGNTIIHVRTDITLHEPMYIF 64
QY      65 LCMISGIDILISTSMPKMLAIFWENSTTIQPDACLLQIPAHISLSEGESVILLAMAADR 124
Db      65 LAMALATDGLCLSTLPVLGI FWFDAEIGIPACFTQLFIHLSLVESVLLSMSEFDR 124
QY      125 YVALCHPLRHATVTLTPRVTKIGVAAYVGAALMAFLPEVFIKQLPFCRSNIISHSYCLHQ 184
Db      125 YVALCHPLRHATVTLTPRVTKIGVAAYVGAALMAFLPEVFIKQLPFCRSNIISHSYCLHQ 184
QY      185 DVMKLACDDIRVVNVYGLIIVISAIGDLSLISFYLILKTVTGLTREACA-KAFETCV 243
Db      185 DVMKLACDDIRVVNVYGLIIVISAIGDLSLISFYLILKTVTGLTREACA-KAFETCV 243
QY      244 SHVCAVFIFYVPFISLWVHRFSKRDSPLPVILANIYLVPLVPLNPIYGVKTKRIOR 303
Db      244 SHVCAVFIFYVPFISLWVHRFSKRDSPLPVILANIYLVPLVPLNPIYGVKTKRIOR 303
QY      304 SHICAVLLFYTPMIGLSVHRFGKAPHLVQVWGMFVLLFPVPMNIIVISVKTQIKRDR 304
Db      304 SHICAVLLFYTPMIGLSVHRFGKAPHLVQVWGMFVLLFPVPMNIIVISVKTQIKRDR 304

```

```

QY      304 ILRLP 308
Db      305 IIRKF 309

```

## RESULT 10

```

Q8VGZ6 PRELIMINARY; PRT; 314 AA.
ID      08VGZ6
AC      01-MAR-2002 (Tremblrel. 20, Created)
DT      01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT      01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE      Olfactory receptor MOR14-2.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN      [1]
RP      NCBI_TaxId=10090;
RA      Zhang X., Firestein S.J.;
RT      "The olfactory receptor gene superfamily of the mouse.";
RL      Nat. Neurosci. 0:0-0(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Adams M.;
RL      Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR      EMBL; AY072994; AAL60657.1; -
DR      InterPro; IPR000276; GPCR_Rhodopsn.
DR      Pfam; PF00001; 7tm_1; 1.
DR      PRINTS; PR00237; GPCR_RHODOPSN.
DR      PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR      PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
KM      Receptor.
SQ      SEQUENCE 314 AA; 35096 MW; 9F1016F881A040E9 CRC64;

```

Query Match 54.1%; Score 875.5; DB 11; Length 314;

Best Local Similarity 52.3%; Pred. No. 7.1e-70; Matches 162; Conservative 63; Mismatches 82; Indels 3; Gaps 2;

```

QY      4 PNGESSA--TFEIIIGPGLBEAOFMLAPLCSYLYLAVGNLTIIYIVTEHSLHEPM 61
Db      2 PSNENSTAPVPVFLTGIGLSTSHTWISIFPCCLYALISGNSMILFVITTESSLHEM 61
QY      62 YIFLCMLSGIDILISTSMPKMLAIFWENSTTIQPDACLLQIPAHISLSEGESVILLAMA 121
Db      62 YIFLCMLSGIDILISTSMPKMLAIFWENSTTIQPDACLLQIPAHISLSEGESVILLAMA 121
QY      122 FDRYVACHPLRHATVTLTPRVTKIGVAAYVGAALMAFLPEVFIKQLPFCRSNIISHSYC 181
Db      122 FDRYVACHPLRHATVTLTPRVTKIGVAAYVGAALMAFLPEVFIKQLPFCRSNIISHSYC 181
QY      182 LHODVMKLACDDIRVVNVYGLIIVISAIGDLSLISFYLILKTVTGLTREACA-KAF 240
Db      182 FHRDVMKLACDDIRVVNVYGLIIVISAIGDLSLISFYLILKTVTGLTREACA-KAF 241
QY      241 TCYSHVCAVFIFYVPFISLWVHRFSKRDSPLPVILANIYLVPLVPLNPIYGVKTKRI 300
Db      242 TCYSHVCAVFIFYVPFISLWVHRFSKRDSPLPVILANIYLVPLVPLNPIYGVKTKRI 301
QY      301 RQRIILRFV 310
Db      302 RKAMLVFV 311

```

## RESULT 11

```

Q8VGZ2 PRELIMINARY; PRT; 314 AA.
ID      08VGZ2
AC      01-MAR-2002 (Tremblrel. 20, Created)
DT      01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT      01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE      Olfactory receptor MOR30-1.
OS      Mus musculus (Mouse).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang X., Firestein S.J.,  
 RT "The olfactory receptor gene superfamily of the mouse.",  
 RL Nat. Neurosci. 0:0-0(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.,  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY072998; AL60661.1;  
 DR InterPro; IPR000276; GPCR\_Rhodopsin.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCRHOOPS.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_FL\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECP\_FL\_2; 1.  
 KM Receptor.  
 SQ SEQUENCE 314 AA; 35577 MW; 718EFEF2F351F669 CRC64;  
 Query Match 54.0%; Score 873.5; DB 11; Length 314;  
 Best Local Similarity 50.2%; Pred. No. 1,1e-69;  
 Matches 155; Conservative 69; Mismatches 85; Indels 1; Gaps 1;  
 QY 1 MVDNGNESSATYFILGLPGLAEAPFLPCLSYLAVGNLTITVIREHSIHP 60  
 DB 1 MVASNSSSHPLFEMLLIGLENYQFWIAFPCCVAVIALGNITLIIIDITLHP 60  
 QY 61 MYFLCMLSGIDILISTSSPKMLATFWNSTTIOPDACLLOIFAIHSLSGMSTVLLAM 120  
 DB 61 MYFLALATLTDVLSSTQPKMLALMHSHEIENACLIQVFIHASSVESGLMTM 120  
 QY 121 AFDRYVAICPLRAVYLTLPVTKIGVAAVVGAALMAPLVEFIKQLPFCRSNLSHSY 180  
 DB 121 ALDRYVAICPLRHSSILITTSVVIKGAAMVVRGLIMVSPFCFMSRMFPCNKVYIPQSY 180  
 QY 181 CLHODVWKLACDDIRVNVVGLVITSAIGLDSLISFSLILKTYLGL-TREAQAKAF 239  
 DB 181 CEMNAVILKVCADTRNKGVLFAVSVGFDTIVSVYWMILRVLHPGEGARLRAF 240  
 QY 240 GTCVSHVCAVIFVPPIGLSMVRFSKRDSPLVILANILVLPVLPNPIYGVYKTE 299  
 DB 241 GTCASHVCVILAFYIALFTFLTRGHVPRVHMFANFPLVPPMLNPIYGVYKTE 300  
 QY 300 IRORIILRF 308  
 DB 301 IRDVRIRGF 309  
 RESULT 12  
 Q8VGZ3 PRELIMINARY; PRT; 318 AA.  
 ID Q8VGZ3  
 AC Q8VGZ3;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Olfactory receptor MOR10-2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang X., Firestein S.J.,  
 RT "The olfactory receptor gene superfamily of the mouse.",  
 RL Nat. Neurosci. 0:0-0(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.,  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY072997; AL60660.1;  
 DR InterPro; IPR001064; Crystallin.

DR InterPro; IPR000276; GPCR\_Rhodopsin.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCRHOOPS.  
 DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECP\_FL\_2; 1.  
 KM Receptor.  
 SQ SEQUENCE 318 AA; 35593 MW; 3A9699B821087E48 CRC64;  
 Query Match 54.0%; Score 873.5; DB 11; Length 318;  
 Best Local Similarity 54.4%; Pred. No. 1,1e-69;  
 Matches 168; Conservative 55; Mismatches 85; Indels 1; Gaps 1;  
 QY 1 MVDNGNESSATYFILGLPGLAEAPFLPCLSYLAVGNLTITVIREHSIHP 60  
 DB 1 MADNHSQSHLYVILGIPGLQKYYMAFPGLGATVIALFQNGVILISTKESSLHIP 60  
 QY 61 MYFLCMLSGIDILISTSSPKMLATFWNSTTIOPDACLLOIFAIHSLSGMSTVLLAM 120  
 DB 61 MYFLCMLAFADMGTLCTLPMLGIFWYKRTTDPGCLVQWYFIHTFSAISGLVAM 120  
 QY 121 AFDRYVAICPLRAVYLTLPVTKIGVAAVVGAALMAPLVEFIKQLPFCRSNLSHSY 180  
 DB 121 ALDRYVAICPLRAVYLTLPVTKIGVAAVVGAALMAPLVEFIKQLPFCRSNLSHSY 180  
 QY 181 CLHODVWKLACDDIRVNVVGLVITSAIGLDSLISFSLILKTYLGL-TREAQAKAF 239  
 DB 181 CLHODVWKLACDDIRVNVVGLVITSAIGLDSLISFSLILKTYLGL-TREAQAKAF 240  
 QY 240 GTCVSHVCAVIFVPPIGLSMVRFSKRDSPLVILANILVLPVLPNPIYGVYKTE 299  
 DB 241 NTCVSHCAVILFVPPIGLSVIRHREKSLPLTHALMAVILVPPVLPNPIYGVYKTE 300  
 QY 300 IRORIILRF 308  
 DB 301 IRKRIIOIF 309  
 RESULT 13  
 Q8VH01 PRELIMINARY; PRT; 317 AA.  
 ID Q8VH01  
 AC Q8VH01;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Olfactory receptor MOR19-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang X., Firestein S.J.,  
 RT "The olfactory receptor gene superfamily of the mouse.",  
 RL Nat. Neurosci. 0:0-0(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.,  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY072987; AL60650.1;  
 DR InterPro; IPR000276; GPCR\_Rhodopsin.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCRHOOPS.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_FL\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECP\_FL\_2; 1.  
 KM Receptor.  
 SQ SEQUENCE 317 AA; 35357 MW; 165EC74EC379CA7 CRC64;  
 Query Match 53.3%; Score 862.5; DB 11; Length 317;  
 Best Local Similarity 53.2%; Pred. No. 1e-68;  
 Matches 158; Conservative 59; Mismatches 79; Indels 1; Gaps 1;  
 QY 15 ILGLPGLAEAPFLPCLSYLAVGNLTITVIREHSIHPMYIFLCMLSGIDIL 74

Db 15 LITGIPGLEDVQWPCIPICWVYIALALNCTILVIRTSLSHPQYIFLSMLATDVG 74  
Qy 75 ISTSNPKKALAFWENSTTIQFDACILQIFAHISJSGMESTVLLAMADRYAALCHPLRH 134  
Db 75 LSVSTLPTLVANFLLNHRDIEFHSCLTQMFIFHTFSMESALILLMAFDRFAVAINSLHY 134  
Qy 135 AVVLTLPRTVKIGVAAVRGALMAPLPVFIKQLPFCRSNIIHSYCCHODVMKACDII 194  
Db 135 TAVLPTTRIKIGLAAVRGVWMLTPIILKLPFCGVIIISHCYCHPDMKACGPV 194  
Qy 195 RVNVVGLIIVISALGDSLLISFSYLLIKTVLGL-TREAQAKAFGTCVSHVCAVFIY 253  
Db 195 RVNIYIGSLVYCSFQVDSVFIIVISYLLIKTVLGIASGDGKLKALNTCVSHIFVFIY 254  
Qy 254 VPFIGSMVHFRFSKRDSPPLVILANIVLVPVNPVIVYGVKTEIRILRLEHV 310  
Db 255 VPLVIALIHRGTFASPLIHTVMANLFLFPLVNPVLSLTKQISAVACKLIFKV 311

## RESULT 14

Q8VGM0 PRELIMINARY; PRT; 319 AA.

AC Q8VGM0  
DT 01-MAR-2002 (TEMBLrel. 20, Created)  
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)  
DE Olfactory receptor MOR30-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN NCB1  
RP SEQUENCE FROM N.A.  
RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse."  
RL Nat. Neurosci. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY073034; AL60697.1; -  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm.1; 1.  
DR PRINTS; PR00237; GPCR\_RHODOPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 319 AA; 35988 MW; B2DB4A1465AA063F CRC64;

Query Match 53.0%; Score 857.5; DB 11; Length 319;  
Best Local Similarity 49.5%; Pred. No. 2,9e-68;  
Matches 154; Conservative 66; Mismatches 90; Indels 1; Gaps 1;

Qy 1 WDPNGNESSATYFLLIGPGLSEAQFWLAPLCSLYLAVGNLTIIYVTEHSLHP 60  
Db 6 MASNSSSVSSTFFLLIGPGLSENYQFWAFPCMYIVAVGNLTIIHRIHTLHP 65  
Qy 61 MYFLCMLSGDILISTSSMPKLAIFWENSTTIQFDACILQIFAHISJSGMESTVLLAM 120  
Db 66 MYFLCMLSTLTDVSSSTQPKLALFVHDKIYHACLQVFFIHFSSVESGVLMTM 125  
Qy 121 AFRVYVAICHLPHATVLTLPRTVKIGVAAVRGALMAPLPVFIKQLPFCRSNIIHSY 180  
Db 126 ALDRYVAICPLPHATVLTLPRTVKIGVAAVRGALMAPLPVFIKQLPFCRSNIIHSY 185  
Qy 181 CHQDVMKACDIDRVNVVYGLIVITSAIGDSLLISFSYLLIKTVLGL-TREAQAKAF 239  
Db 186 CEHMAVILKVCADTRVNRGIGLFTVAFSVGPDIIIVISYVWILDAVRLSGERLKA 245  
Qy 240 GTCVSHVCAVFIYVFFIGLSMVRFSKRDSPPLVILANIVLVPVNPVIVYGVKTE 299  
Db 246 GTCASHIGVILTLTYPALFTPLTHFGHHVPRVHIMFANVYLLVPPMLNPIIGVTRTK 305

Qy 300 IRORIILRLEHV 310  
Db 306 IRORYTRGFCV 316

## RESULT 15

Q8VGM0 PRELIMINARY; PRT; 319 AA.

AC Q8VGM0  
DT 01-MAR-2002 (TEMBLrel. 20, Created)  
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)  
DE Olfactory receptor MOR17-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN NCB1  
RP SEQUENCE FROM N.A.  
RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse."  
RL Nat. Neurosci. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY072992; AL60655.1; -  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm.1; 1.  
DR PRINTS; PR00237; GPCR\_RHODOPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 319 AA; 36040 MW; D3FP2D8755FDC2F CRC64;

Query Match 53.0%; Score 857.5; DB 11; Length 319;  
Best Local Similarity 52.9%; Pred. No. 2,9e-68;  
Matches 163; Conservative 57; Mismatches 87; Indels 1; Gaps 1;

Qy 1 WDPNGNESSATYFLLIGPGLSEAQFWLAPLCSLYLAVGNLTIIYVTEHSLHP 60  
Db 1 MASNSVSVSSTFFLLIGPGLSENYQFWAFPCMYIVAVGNLTIIHRIHTLHP 60  
Qy 61 MYFLCMLSGDILISTSSMPKLAIFWENSTTIQFDACILQIFAHISJSGMESTVLLAM 120  
Db 61 MYFLCMLSTLTDVSSSTQPKLALFVHDKIYHACLQVFFIHFSSVESGVLMTM 120  
Qy 121 AFRVYVAICHLPHATVLTLPRTVKIGVAAVRGALMAPLPVFIKQLPFCRSNIIHSY 180  
Db 121 ALDRYVAICPLPHATVLTLPRTVKIGVAAVRGALMAPLPVFIKQLPFCRSNIIHSY 180  
Qy 181 CHQDVMKACDIDRVNVVYGLIVITSAIGDSLLISFSYLLIKTVLGL-TREAQAKAF 239  
Db 181 CHQDIIIRLACADIRNVVYGLIVITSAIGDSLLISFSYLLIKTVLGL-TREAQAKAF 240  
Qy 240 GTCVSHVCAVFIYVFFIGLSMVRFSKRDSPPLVILANIVLVPVNPVIVYGVKTE 299  
Db 241 NTGASHICAVLILIVYMIGLSIYHRAKSSPRVHIFMAHITLWPPVNPPIIYSKTK 300  
Qy 300 IRORIILR 307  
Db 301 IROGIPLH 308

Search completed: February 9, 2004, 16:25:11  
Job time: 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 16:18:01 ; Search time 17 Seconds  
(without alignments)

876,910 Million cell updates/sec

Title: US-10-017-066a-2

Perfect score: 1617  
Sequence: 1 MVDPNNESSATYFTLLGLP.....KEIRQRLRLPHVATHASEP 317

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	979	60.5	320	1 OX82_HUMAN	Q9H255 homo sapien
2	977	60.4	320	1 OX82_RAT	Q9H255 ratius norv
3	868.5	53.7	312	1 OX12_HUMAN	Q9H344 homo sapien
4	806.5	49.9	314	1 OX12_HUMAN	Q9H344 homo sapien
5	792.5	49.0	318	1 OX12_HUMAN	Q9H346 homo sapien
6	774.5	47.9	315	1 OX12_HUMAN	Q9H341 homo sapien
7	762	47.1	317	1 OX12_HUMAN	Q9H342 homo sapien
8	751	46.4	312	1 OX12_HUMAN	Q9H342 homo sapien
9	728.5	45.1	312	1 OX12_HUMAN	Q9H342 homo sapien
10	704	43.5	313	1 OX12_HUMAN	Q9H342 homo sapien
11	632	39.1	310	1 OX12_HUMAN	Q9H342 homo sapien
12	464	28.7	319	1 OX12_HUMAN	Q9H342 homo sapien
13	439	27.1	312	1 OX12_HUMAN	Q9H342 homo sapien
14	434.5	26.9	312	1 OX12_HUMAN	Q9H342 homo sapien
15	432.5	26.7	312	1 OX12_HUMAN	Q9H342 homo sapien
16	431.5	26.7	310	1 OX12_HUMAN	Q9H342 homo sapien
17	431.5	26.7	314	1 OX12_HUMAN	Q9H342 homo sapien
18	430.5	26.6	319	1 OX12_HUMAN	Q9H342 homo sapien
19	428.5	26.5	311	1 OX12_HUMAN	Q9H342 homo sapien
20	428	26.5	317	1 OX12_HUMAN	Q9H342 homo sapien
21	427.5	26.4	314	1 OX12_HUMAN	Q9H342 homo sapien
22	427	26.4	323	1 OX12_HUMAN	Q9H342 homo sapien
23	426	26.3	315	1 OX12_HUMAN	Q9H342 homo sapien
24	425.5	26.3	314	1 OX12_HUMAN	Q9H342 homo sapien
25	425.5	26.3	320	1 OX12_HUMAN	Q9H342 homo sapien
26	424.5	26.3	315	1 OX12_HUMAN	Q9H342 homo sapien
27	423.5	26.2	312	1 OX12_HUMAN	Q9H342 homo sapien
28	423.5	26.2	321	1 OX12_HUMAN	Q9H342 homo sapien
29	421	26.0	317	1 OX12_HUMAN	Q9H342 homo sapien
30	421	26.0	317	1 OX12_HUMAN	Q9H342 homo sapien
31	420.5	26.0	312	1 OX12_HUMAN	Q9H342 homo sapien
32	420.5	26.0	313	1 OX12_HUMAN	Q9H342 homo sapien
33	418	25.9	311	1 OX12_HUMAN	Q9H342 homo sapien

## ALIGNMENTS

34	416.5	25.8	314	1 OX12_HUMAN	P58181 homo sapien
35	415.5	25.7	312	1 OX12_HUMAN	P37070 gallus galli
36	413.5	25.6	311	1 OX12_HUMAN	O76001 homo sapien
37	413	25.5	312	1 OX12_HUMAN	P34982 homo sapien
38	412.5	25.5	318	1 OX12_HUMAN	P37067 gallus galli
39	411.5	25.4	311	1 OX12_HUMAN	O15620 homo sapien
40	411	25.4	311	1 OX12_HUMAN	O95154 canis fami
41	409.5	25.3	313	1 OX12_HUMAN	P58173 homo sapien
42	409.5	25.3	313	1 OX12_HUMAN	P23265 ratius norv
43	407.5	25.2	318	1 OX12_HUMAN	P37068 gallus galli
44	407	25.2	313	1 OX12_HUMAN	P30955 canis fami
45	407	25.2	318	1 OX12_HUMAN	Q9Y449 homo sapien

RESULT 1  
OX12\_HUMAN STANDARD; PRT; 320 AA.  
ID OX12\_HUMAN  
AC Q9H255;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Olfactory receptor 51b2 (prostate specific G-protein coupled receptor) (HPRAD).  
GN OR51B2 OR PSGR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=21002538; PubMed=1118034;  
RX Xu L.L., Stackhouse B.G., Florence K., Zhang W., Shanmugan N., Sesterhenn I.A., Zou Z., Srikantan V., Augustus M., Roach S., Carter K., McLeod D.G., Mou J.W., Sopper D., Srivastava S., "PSGR, a novel prostate-specific gene with homology to a G protein-coupled receptor, is overexpressed in prostate cancer."; Cancer Res. 60:6568-6572(2000).  
RL  
RN [2]  
RP MEDLINE=21564169; PubMed=11707321;  
RX Yuan T.T., Toy P., McLary J.A., Lin R.J., Miyamoto N.G., Kretschmer P.J.; "Identification of a prostate-specific G-protein coupled receptor (PSGR) that interacts with G alpha subunit in yeast two hybrid assay."; Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RL  
RN [3]  
RP MEDLINE=21564169; PubMed=11707321;  
RX Yuan T.T., Toy P., McLary J.A., Lin R.J., Miyamoto N.G., Kretschmer P.J.; "Cloning and genetic characterization of an evolutionarily conserved human olfactory receptor that is differentially expressed across species."; Gene 278:41-51(2001).  
RL  
RN [4]  
RP MEDLINE=22388257; PubMed=12477932;  
RX Straube R.L., Feingold E.A., Grouse L.H., Derge J.G., Krausner R.D., Collins F.S., Wagner L., Schmeier C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedlin T.B., Toshiyuki S., Carinci P., Prange C., Raba S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vallatton D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,



RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,  
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Maitra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences",  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXCLUSIVELY EXPRESSED IN THE PROSTATE.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF311306; AAG40776.1; -  
 CC EMBL; AF369708; AAK38728.1; -  
 CC EMBL; AY033942; AAK57550.1; -  
 CC EMBL; BC020768; AAK20758.1; -  
 CC Genew; HGNC:15195; OR51E2.  
 DR GO; GO:0016021; C: integral to membrane; NAS.  
 DR GO; GO:0004984; F: olfactory receptor activity; NAS.  
 DR GO; GO:0007608; F: olfactory; NAS.  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PRO0237; GPCR\_RHODOPSIN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KM G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 23  
 FT TRANSSEM 1 48  
 FT DOMAIN 49 59  
 FT TRANSSEM 60 84  
 FT DOMAIN 85 100  
 FT TRANSSEM 101 119  
 FT DOMAIN 120 142  
 FT TRANSSEM 143 162  
 FT DOMAIN 163 197  
 FT TRANSSEM 198 221  
 FT DOMAIN 222 235  
 FT TRANSSEM 236 260  
 FT DOMAIN 261 292  
 FT TRANSSEM 293 320  
 FT DOMAIN 320 320  
 FT DISULFID 5  
 FT CARBOHYD 5  
 SQ SEQUENCE 320 AA; 35492 MW; 03582CC2AAB6E2C6 CRC64;  
 Query Match 60.5%; Score 979; DB 1; Length 320;  
 Best Local Similarity 59.5%; Pred. No. 1,4e-68;  
 Matches 182; Conservative 55; Mismatches 67; Indels 2; Gaps 2;

QY 246 VCAVFIFVFPFGISMTWRFSGKRDSPPLVLIANTLIVPVLNPIYGVTKKEIKRIL 305  
 Db 244 IGVVLAFFVPLGLSGVHRFPNGSLHPVIVVMGDYLLPVPINPIYGAATKQIRIVL 303  
 QY 306 RLFFVA 311  
 Db 304 AMFKIS 309  
 RESULT 2  
 OXK2\_RAT  
 ID OXK2\_RAT STANDARD; PRT; 320 AA.  
 AC 088628;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Olfactory receptor 51E2 (G-protein coupled receptor RALC).  
 GN OR51E2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OC NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSU=Brain;  
 RC MEDLINE=99131082; PubMed=9932290;  
 RA Raming K., Konzelmann S., Breer H.;  
 RT "Identification of a novel G-protein coupled receptor expressed in  
 RT distinct brain regions and a defined olfactory zone";  
 RL Recept. Channels 6:141-151(1998).  
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN SOME AREAS OF THE BRAIN AND  
 CC IN THE OLFACTORY EPITHELIUM.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF079864; AAD12761.1; -  
 CC InterPro; IPR00276; GPCR\_Rhodopsn.  
 CC Pfam; PF00001; 7tm.1; 1.  
 CC PRINTS; PRO0237; GPCR\_RHODOPSIN.  
 CC PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 CC PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KM G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 23  
 FT TRANSSEM 24 48  
 FT DOMAIN 49 59  
 FT TRANSSEM 60 84  
 FT DOMAIN 85 94  
 FT TRANSSEM 95 119  
 FT DOMAIN 120 138  
 FT TRANSSEM 139 162  
 FT DOMAIN 163 197  
 FT TRANSSEM 198 221  
 FT DOMAIN 222 235  
 FT TRANSSEM 236 260  
 FT DOMAIN 261 292  
 FT TRANSSEM 293 320  
 FT DOMAIN 320 320  
 FT DISULFID 5  
 FT CARBOHYD 5  
 SQ SEQUENCE 320 AA; 35505 MW; E7FF78F5F5B9F4 CRC64;  
 Query Match 60.4%; Score 977; DB 1; Length 320;  
 Best Local Similarity 59.5%; Pred. No. 2e-68;  
 Matches 182; Conservative 53; Mismatches 69; Indels 2; Gaps 2;

QY 7 NESSATYFLLIGLPGLEENQFWLAPPLCSLYIANVGNLTIIYVRETHSLHEPMYFLC 66  
 DB 5 NETHAT-FMLIGIPGLEENHFWGFPFLSMYAVAFGCIIVFIRKTRSRSHAPYFLC 63  
 QY 67 MMSGIDILISTSSMPKMLAFWENSTTIOFDACLIQIFAHSLSGSESTVLLAMAFDRYV 126  
 DB 64 MMLADIDALSTSTWPKLALFWDSREITFDACLAGMFHLSGSESTVLLAMAFDRYV 123  
 QY 127 AICHPRLHATVTLPRVTKIGVAAYVRGALMAPLPVFIKOLPFCRSNLSHSYCLHODV 186  
 DB 124 AICHPRLHAAVNTVQVIGVAVLVRGSLFFPPLIKRLAFCHSNVLSHSYCVHODV 183  
 QY 187 MMLACDIDIRVNVYGLIVISAIGLSLISFSYLLIKTVLGL-TREAAKAFGTCVSH 245  
 DB 184 MMLAVDTIPNVYGLITALLVWGVDMFISLSTYLLIRAVQLSKSKERAKAFGTCVSH 243  
 QY 246 VCAVFIYVPTGLSMVHRFSKRSDPLFVILANTYLLVPVLPVNPVYGVKIKERQRL 305  
 DB 244 IGVVLAFFYPLIGLSVHRFGNSLDPIYHVMGDVYLLPVIINDIYGAATKQIRTVL 303  
 QY 306 RLEHVA 311  
 DB 304 AMFKIS 309

RESULT 3  
 OX12\_HUMAN STANDARD; PRT; 312 AA.  
 ID OX12\_HUMAN  
 AC Q9H344;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Olfactory receptor 5112 (HORS/Beta12).  
 GN OR5112.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20570519; PubMed=11121057;  
 RA Bulger M., Bender M.A., van Doorninck J.H., Wertman B., Farrell C.M.,  
 RA Felsenfeld G., Groudine M., Hardison R.;  
 RT "Comparative structural and functional analysis of the olfactory  
 RT receptor genes flanking the human and mouse beta-globin gene  
 RT clusters.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 97:14560-14565(2000).  
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/  
 CC or send an email to license@isb-sdb.ch).  
 CC  
 DR EMBL; AF137396; AAG41678.1; -;  
 DR GO; GO:0016021; C:integral to membrane; NAS.  
 DR GO; GO:0004984; P:olfactory receptor activity; NAS.  
 DR GO; GO:0007608; P:olfactory receptor activity; NAS.  
 DR InterPro; IPR000276; GPCR\_Rhodopsin.  
 DR Pfam; PF00001; 7tm\_1.1.  
 DR PRINTS; PR00237; GPCRHOPOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECP\_F1\_2; 1.  
 KM G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KM Multigene family; Olfaction.  
 FT DOMAIN 1 25 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 26 49 1 (POTENTIAL).

FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 79 2 (POTENTIAL).  
 FT DOMAIN 80 100 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 101 120 3 (POTENTIAL).  
 FT DOMAIN 121 139 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 140 158 4 (POTENTIAL).  
 FT DOMAIN 159 195 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 196 219 5 (POTENTIAL).  
 FT DOMAIN 220 236 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 237 259 6 (POTENTIAL).  
 FT DOMAIN 260 274 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 275 294 7 (POTENTIAL).  
 FT DOMAIN 295 312 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 97 189 BY SIMILARITY.  
 FT CARBOHYD 5 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 312 AA; 35002 MW; F17CD16C95FD1AF6 CRC64;

Query Match 53.7%; Score 868.5; DB 1; Length 312;  
 Best Local Similarity 53.6%; Pred. No. 4.5e-60;  
 Matches 163; Conservative 51; Mismatches 89; Indels 1; Gaps 1;

QY 7 NESSATYFLLIGLPGLEENQFWLAPPLCSLYIANVGNLTIIYVRETHSLHEPMYFLC 66  
 DB 5 NETHAT-FMLIGIPGLEENHFWGFPFLSMYAVAFGCIIVFIRKTRSRSHAPYFLC 63  
 QY 67 MMSGIDILISTSSMPKMLAFWENSTTIOFDACLIQIFAHSLSGSESTVLLAMAFDRYV 126  
 DB 64 MMLADIDALSTSTWPKLALFWDSREITFDACLAGMFHLSGSESTVLLAMAFDRYV 123  
 QY 127 AICHPRLHATVTLPRVTKIGVAAYVRGALMAPLPVFIKOLPFCRSNLSHSYCLHODV 186  
 DB 124 AICHPRLHAAVNTVQVIGVAVLVRGSLFFPPLIKRLAFCHSNVLSHSYCVHODV 183  
 QY 187 MMLACDIDIRVNVYGLIVISAIGLSLISFSYLLIKTVLGL-TREAAKAFGTCVSH 245  
 DB 184 MMLAVDTIPNVYGLITALLVWGVDMFISLSTYLLIRAVQLSKSKERAKAFGTCVSH 243  
 QY 246 VCAVFIYVPTGLSMVHRFSKRSDPLFVILANTYLLVPVLPVNPVYGVKIKERQRL 305  
 DB 244 IGVVLAFFYPLIGLSVHRFGNSLDPIYHVMGDVYLLPVIINDIYGAATKQIRTVL 303  
 QY 306 RLEHVA 311  
 DB 304 AMFKIS 309

RESULT 4  
 OX11\_HUMAN STANDARD; PRT; 314 AA.  
 ID OX11\_HUMAN  
 AC Q9H343;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Olfactory receptor 5111 (HORS/Beta11).  
 GN OR5111.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20570519; PubMed=11121057;  
 RA Bulger M., Bender M.A., van Doorninck J.H., Wertman B., Farrell C.M.,  
 RA Felsenfeld G., Groudine M., Hardison R.;  
 RT "Comparative structural and functional analysis of the olfactory  
 RT receptor genes flanking the human and mouse beta-globin gene  
 RT clusters.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 97:14560-14565(2000).  
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
 CC EMBL; AF137396; AAG1679.1; -  
 DR GO; GO:0016021; C:integral to membrane; NAS.  
 DR GO; GO:0004984; F:olfactory receptor activity; NAS.  
 DR GO; GO:0007608; P:olfaction; NAS.  
 DR InterPro: IPR000276; GPCR\_Rhodopsin.  
 DR Pfam: PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCR\_Rhodopsin.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G-PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane;  
 KW Multigene family; Olfaction.  
 FT DOMAIN 1 27  
 FT TRANSMEM 25 51  
 FT DOMAIN 52 59  
 FT TRANSMEM 60 81  
 FT DOMAIN 82 102  
 FT TRANSMEM 103 122  
 FT DOMAIN 123 141  
 FT TRANSMEM 142 160  
 FT DOMAIN 161 197  
 FT TRANSMEM 198 221  
 FT DOMAIN 222 238  
 FT TRANSMEM 239 261  
 FT DOMAIN 262 276  
 FT TRANSMEM 277 296  
 FT DOMAIN 297 314  
 FT DISULFID 99 191  
 SQ SEQUENCE 314 AA; 35483 MW; 08B40AD79CB94911 CRC64;

Query Match 49.9%; Score 806.5; DB 1; Length 314;  
 Best Local Similarity 48.4%; Pred. No. 2.7e-55;  
 Matches 150; Conservative 66; Mismatches 93; Indels 1; Gaps 1;

QY 1 MVDNNGESSATYFLLIGLPGLEBAQFWLAFLCSLYLAIVGNLTIIYVTHSHHP 60  
 DB 1 MGLNGPFPAPLTLQTLGPIQTLGTLWALLFCILYMSIVGNLSITLVWEPLHOP 60  
 QY 61 MYFLCMTSGIDILISTSNPKMAIFWENSTTIQPDACLIQFAIHSLSGNESTVLAM 120  
 DB 61 MYFLSKMALNDLGVSPSTLPVISTFCFVYHVAFNACLVQMFHTFSFVESGITLAM 120  
 QY 121 AFDRYVALCHPLRHATVLTLPRTKIGVAAYVGAALMAPLPVFQQLPFGCSNLSHSY 180  
 DB 121 SLDFEVALCPPLVYVLTNRLIANGGLITLSFTLLPFPVVRPLPFGKGNVLHSHY 180  
 QY 181 CLHQDVAKLACDDIRVNVVYGLIIVISAIGLSDLLISFSYLLIKTLVGL-TREAQAKAF 239  
 DB 181 CLHPDLAKVACGDIHVNNYIGLVLIIFTYGMDSTFLLISYALLIRKMLVITISQEQRLKAL 240  
 QY 240 GTGVSHVCAVFITYVPIGSMWHRFSKRKDSPLPILANTYILVPPVLANPVIYGYKYE 299  
 DB 241 NTCSHSCAVLAAYVPLIAVSMIRFWKSPAPVHVHMSNVYLFVPMNLPIIYSVKYE 300  
 QY 300 IRORIILRFH 309  
 DB 301 IRKGIILKFFH 310  
 RESULT 5  
 ID OYD1 HUMAN STANDARD; PRT; 318 AA.  
 AC O9H346;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Olfactory receptor 52D1 (HORS'beta14).

GN OR52D1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=20570519; PubMed=11121057;  
 RA Bulger M., Bender M.A., van Doornink J.H., Wertman B., Farrell C.M.,  
 RA Felsenfeld G., Grondine M., Hardison R.;  
 RT "Comparative structural and functional analysis of the olfactory  
 RT receptor genes flanking the human and mouse beta-globin gene  
 RT clusters";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14560-14565(2000).  
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF137396; AAG1676.1; -  
 DR GO; GO:0016021; C:integral to membrane; NAS.  
 DR GO; GO:0004984; F:olfactory receptor activity; NAS.  
 DR GO; GO:0007608; P:olfaction; NAS.  
 DR InterPro: IPR000276; GPCR\_Rhodopsin.  
 DR Pfam: PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCR\_Rhodopsin.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G-PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction.  
 FT DOMAIN 1 28  
 FT TRANSMEM 29 52  
 FT DOMAIN 53 60  
 FT TRANSMEM 61 82  
 FT DOMAIN 83 103  
 FT TRANSMEM 104 123  
 FT DOMAIN 124 142  
 FT TRANSMEM 143 161  
 FT DOMAIN 162 198  
 FT TRANSMEM 199 222  
 FT DOMAIN 223 239  
 FT TRANSMEM 240 262  
 FT DOMAIN 263 278  
 FT TRANSMEM 279 298  
 FT DOMAIN 299 318  
 FT DISULFID 100 192  
 FT CARBOHYD 5 5  
 SQ SEQUENCE 318 AA; 35121 MW; 178968087D613CB2 CRC64;

Query Match 49.0%; Score 792.5; DB 1; Length 318;  
 Best Local Similarity 46.2%; Pred. No. 3.3e-54;  
 Matches 145; Conservative 71; Mismatches 93; Indels 5; Gaps 4;

QY 1 MVDN--GNSSATYFLLIGLPGLEBAQFWLAFLCSLYLAIVGNLTIIYVTHSHH 58  
 DB 1 MSDNLSDNHLPDFF-LTGLPGLEBAHFIAPFCMYIVLVAGNALILVIAAMDALH 59  
 QY 59 EPMYFLCMTSGIDILISTSNPKMAIFWENSTTIQPDACLIQFAIHSLSGNESTVL 118  
 DB 60 APMYFLCMTSLTDIALSTTPVPMALILWLHAGEISFGGLAQMPFVSHIYALSSILL 119  
 QY 119 AMAEDRYVALCHPLRHATVLTLPRTKIGVAAYVGAALMAPLPVFQQLPFGCSNLSH 178  
 DB 120 AMAEDRYVALCHPLRITLIHNAVIGVGLFRSAIYSPFIPLRLPYGCHRMTH 179  
 QY 179 SYCHQDVAKLACDDIRVNVVYGLIIVISAIGLSDLLISFSYLLIKTLVGL-TREAQAK 237

Dd		180	TTCENHGIAIRLACANTTNYIVGLFVALLAMGLDSIIILAIISYGFILNAVHFLDPSHDQKH	239
Oy		238	AFGCASHVAQAEIFVPVPIGISMWHRSEKRR-DSPILVTLLANTYLVLPVNDIYGAK	296
Dd		240	ALSTCGSHICIIIVFIIPAPSFPLTRFGHEHVEPVKATVFILANKVLVLPVNPILYGAR	299
Oy		297	TXEIRQRILLRFFHV 310      : : : :	
Dd		300	TKEIRSRLLIKLTLHL 313	
RESULT 6				
ID	OXML_HUMAN	STANDARD,	PRT,	315 AA.
AC	Q9H341;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Olfactory receptor 51M1 (HOM5.Delta7).			
GN	OR51M1			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxId=9606;			
RX	[1]			
RA	SEQUENCE FROM N.A. MEDLINE=20570519; PubMed=1121057;			
RA	Bulger M., Bender M.A., van Doorninck J.H., Wertman B., Farrell C.M., Feisenfeld G., Grodin M., Hardison R.;			
RT	"Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:14560-14565(2000).			
CC	-1- FUNCTION: PUTATIVE ODORANT RECEPTOR.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation, the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isdb-sib.ch/announce/ or send an email to license@isdb-sib.ch).			
CC	-----			
DR	EMBL; AF13796; AAC41681.1; ALT_INIT.			
DR	GO; GO:0016021; C:integral to membrane; NAS.			
DR	GO; GO:004984; P:olfactory receptor activity; NAS.			
DR	GO; GO:0007608; P:olfactory receptor activity; NAS.			
DR	InterPro; IPR000276; GPCK_Rhodopsn.			
DR	Pfam; PF00001; 7tm_1; 1.			
DR	PRINTS; PR00237; GPCRHOPOBN.			
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.			
DR	PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Multigene family; Olfaction.			
FT	DOMAIN	1	26	EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	27	50	1 (POTENTIAL).
FT	DOMAIN	51	58	CYTOSOLASMIC (POTENTIAL).
FT	TRANSSEM	59	80	2 (POTENTIAL).
FT	DOMAIN	81	101	EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	102	121	3 (POTENTIAL).
FT	DOMAIN	122	140	CYTOSOLASMIC (POTENTIAL).
FT	TRANSSEM	141	159	4 (POTENTIAL).
FT	DOMAIN	160	196	EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	197	220	5 (POTENTIAL).
FT	DOMAIN	221	237	CYTOSOLASMIC (POTENTIAL).
FT	TRANSSEM	238	260	6 (POTENTIAL).
FT	DOMAIN	261	275	EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	276	295	7 (POTENTIAL).
FT	DOMAIN	296	315	CYTOSOLASMIC (POTENTIAL).
FT	TRANSSEM	298	315	BY SIMILARITY.
FT	DISULFID	98	190	

BT	CABOHND	5	5	N-LINKED (GLNAC..)	(POTENTIAL).
SQ	SEQUENCE	315 AA;	35462 MW;	DC192DF5456D2005 CRC64;	
Query Match		47.9%;	Score 774.5;	DB 1;	Length 315;
Best Local Similarity		47.5%;	Pred. No. 7.9e-53;		
Matches 140:	Conservative	65;	Mismatches 89;	Indels 1;	Gaps 1;
OY	14	FILIGLGLLEEAQWLAFLPCSTLYLVNLGULTIIVRTETSHLPEMYIFLCMISGDI	73		
Db	13	FYLSPFGLEGKIKMIPIPFPPMMVAISGSCFILLIKINPRHRTPMYYLSLALTDL	72		
OY	74	LIISSNPKKMLAIWFNSTTIQPACLLQIPAHISLSGMESTVLLANAFDRYVAICHEL	133		
Db	73	GLCVSLPTPTMGIFWENSOIYPFACCOIMFCIHSEFMESVLLMMSFDRFVAICHLR	132		
OY	134	HATVLTLPRTYKIGVAAVVRGAAMAPLPEFIKOLPCRSNIISHSYCLHDVMKLACDD	193		
Db	133	YSVITITQQVVRALYIFRGPAVTITVLTKAFPVGSVLSSHCHGEVQLACTD	192		
OY	194	IRNAVVTGGIYIISAIGLDLSLIFSFLILIKTVLG-TREAQAAGTCYSHVCAYFIF	252		
Db	193	TTFNNILGLMWAVVATVWLDELVALISGLLIHTVAGLASGEORRAQTCAHCAVLVF	252		
OY	253	VYFPGISVMHFESKRDSPLPVIANIYLVLPVNLIYGVGTKEIKRCRIRL	307		
Db	253	FVPWMGI SLVHRFEKHAPPAIHLMNAVYLEFVPMANPIIYSIKTKIHRIAILDL	307		
RESULT 7					
OYB2_HUMAN	OYB2_HUMAN	STANDARD;	PRT;	317 AA.	
ID	OYB2_HUMAN	STANDARD;	PRT;	317 AA.	
AC	OYB2_HUMAN	STANDARD;	PRT;	317 AA.	
DT	28-FEB-2003	(Rel. 41, Created)			
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
D2	Olfactory receptor 52B2.				
GN	OR52B2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE OF 70-287 FROM N.A.				
RA	Fuchs T., Malceva B., Lindhart C., Sharan R., Kuen M., Heizig R.,				
RA	Smullevich D., Elkon R., Steinbach M., O'Brien J.K., Radelof U.,				
RA	Lehmach H., Olander Z., Guesman G., Lancelot D., Shamir R.,				
RT	"DEFOO: a practical scheme for deciphering families of genes."				
RT	Submitted (JUL-2001) to the EMBL/Genebank/DBJ databases.				
CC	-1- FUNCTION: POTATIVE ODORANT RECEPTOR.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
CC	This SWISS-Prot entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; AF399505; AAK9490.1;				
DR	InterPro; IPR000276; GPCR_Rhodopsin.				
DR	Pfam; PF00001; 7tm_1; 1.				
DR	PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.				
DR	PROSITE; PSS0262; G_PROTEIN_RECP_F1_2; 1.				
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;				
KW	Multigene family; Olfaction.				
FT	DOMAIN	1	27		EXTRACELLULAR (POTENTIAL).
FT	TRANSERM	28	51		1 (POTENTIAL).
FT	DOMAIN	52	59		CYTOPASMIC (POTENTIAL).
FT	TRANSERM	60	81		2 (POTENTIAL).
FT	DOMAIN	82	102		EXTRACELLULAR (POTENTIAL).
FT	TRANSERM	103	122		3 (POTENTIAL).

```

FT DOMAIN 123 141 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 142 160 4 (POTENTIAL).
FT DOMAIN 161 197 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 198 221 5 (POTENTIAL).
FT DOMAIN 222 238 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 239 261 6 (POTENTIAL).
FT DOMAIN 262 276 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 277 296 7 (POTENTIAL).
FT DOMAIN 297 317 CYTOPLASMIC (POTENTIAL).
FT DISULFID 99 191 BY SIMILARITY.
FT CARBOHYD 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 317 AA; 35642 MW; DB17DB8791F85E21 CRC64;

Query Match 47.1%; Score 762; DB 1; Length 317;
Best Local Similarity 49.2%; Pred. No. 7.3e-52;
Matches 148; Conservative 57; Mismatches 94; Indels 2; Gaps 2;

QY 14 FILIGLEAEAGFWLAFPLGSLYLAVLGNLTIIYVTEHSHPEPTIFLCMLSGIDI 73
D 14 FVLPGIGLEAHYHWSIPCLTYIRAVGNSILYIVWERTLHPMTFFLSMLAVMDI 73
QY 74 LISTSSMPKALAFWENSTTIQDACLQIFAIHSLSGESTVLLAMAFDRYVAICHPUR 133
D 74 LLSITVPEKALAFWQAHNIAPDACYQGFVHMFVGSATILAMAFDRFVAICAPUR 133
QY 134 HATVLTLPRTKIGVAAYVARGALMAPLPVFIFKQLPFCRSNIIHSYCLHODVMKLACDD 193
D 134 YTVLTLPVTVGRITAVITSPFCIIFPVIFLKLPLFCITNIVPSYCEHIGVARIACAD 193
QY 194 IRVNVVGLIYIISAGLSLISFSYLLIKTVLGL-TREAQAKAFGTCSHVCAVFI 252
D 194 ITVNIWGFSEVPVIMVYLLDVLIVAVSYSLIRAVFPLPSQDARHKALSTGSHLCVILMF 253
QY 253 YVPEIGLSMVRFSKRDSPILVILANIVLVPVLPVYGVKTEIQRLL-ELFHA 311
D 254 YVSPFTLLTHHGKRNIPQHVHILANLVAVPEMLPIYGVKIKQREGVAFHFPDIK 313
QY 312 T 312
D 314 T 314

RESULT 8
OYAL_HUMAN STANDARD; PRT; 312 AA.
AC QYUKL2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Olfactory receptor 52A1 (HPRH10R) (HOK3'Dect4).
GN OS32AL.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9943866; Pubmed=10512676;
RA Ferguson E.A., Penny L.A., Nienhuis A.W., Forget B.G.;
RT "An olfactory receptor gene is located in the extended human beta-
RT globin gene cluster and is expressed in erythroid cells.";
RL Genomics 61:15-23 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20570519; Pubmed=11121057;
RA Balzer M., Bender M.A., van Doorninck J.H., Werrman B., Farrell C.M.,
RA Felsenfeld G., Groudine W., Hardison R.;
RT "Comparative structural and functional analysis of the olfactory
RT receptor genes flanking the human and mouse beta-globin gene
RT clusters.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14560-14565 (2000).
CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

```

```

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: AF154673; AAD51279.1; -
DR EMBL: AF289204; AAG42367.1; -
DR GO: GO:0005887; C: integral to plasma membrane; TAS.
DR GO: GO:0007608; P: olfaction; TAS.
DR GO: GO:0007165; P: signal transduction; TAS.
DR InterPro: IPR00276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCR_Rhodpsn.
DR PROSITE: PS00237; G-PROTEIN RECP_F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN RECP_F1_2; 1.
DR KW G-protein coupled receptor; Transmembrane; Glycoprotein;
DR MultiGene family; Olfaction.
FT DOMAIN 1 27 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 28 51 1 (POTENTIAL).
FT DOMAIN 52 59 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 60 81 2 (POTENTIAL).
FT DOMAIN 82 102 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 103 122 3 (POTENTIAL).
FT DOMAIN 123 141 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 142 160 4 (POTENTIAL).
FT DOMAIN 161 198 5 (POTENTIAL).
FT TRANSSEM 199 222 6 (POTENTIAL).
FT DOMAIN 223 239 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 240 262 6 (POTENTIAL).
FT DOMAIN 263 277 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 278 297 7 (POTENTIAL).
FT DOMAIN 298 312 CYTOPLASMIC (POTENTIAL).
FT DISULFID 99 182 BY SIMILARITY.
FT CARBOHYD 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 312 AA; 35332 MW; BA93279352D513C CRC64;

Query Match 46.4%; Score 751; DB 1; Length 312;
Best Local Similarity 47.8%; Pred. No. 5.1e-51;
Matches 141; Conservative 58; Mismatches 94; Indels 2; Gaps 2;

QY 16 LIGLPGLEAEAGFWLAFPLGSLYLAVLGNLTIIYVTEHSHPEPTIFLCMLSGIDI 75
D 16 LVGIPGLESVQWGIIFCAIYILAMGNSLLISKESSTHEPLIYIFGMGARDIAL 75
QY 76 STSSMPKALAFWENSTTIQDACLQIFAIHSLSGESTVLLAMAFDRYVAICHPUR 135
D 76 ASSIMPOLGIFWENVEIYFSDCLQWFIHTLQGISLVLAMAFDRYVAICYPUR 135
QY 136 TVLFLPRTKIGVAAYVARGALMAPLPVPFK-CLPFCRSNIIHSYCLHODVMKLACDDI 194
D 136 NIFTHQVLTQIGMVVLRALILVAPCLVILKCFQPHRTHTVISHSYCEHAIVKLAANY 195
QY 195 RVNVVGLIYIISAGLSLISFSYLLIKTVLGL-TREAQAKAFGTCSHVCAVFI 253
D 196 QVNIYGLFAVFAVAGDLPFTILSYQIFITFVRFLPQKARPKAFVTCIAHICVPLQFY 255
QY 254 VPEIGLSMVRFSKRDSPILVILANIVLVPVLPVYGVKTEIQRLL-ELF 308
D 256 LIAFSPFTTHRFGSHIPPYTHILPSIYLVAPPPLNPLVGAATTOIRIVAVKMF 310

RESULT 9
OXB2_HUMAN STANDARD; PRT; 312 AA.
ID OXB2_HUMAN
AC Q9Y5P1; O96RD4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update).

```

DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Olfactory receptor 51B2 (HORS/Beta2).  
 GN OR51B2.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP MEDLINE=9238494; PubMed=10220430;  
 RX Bulger M., van Doorninck J.H., Saitoh N., Telling A., Farrell C.,  
 RA Bulger M.A., Felsenfeld G., Axel R., Groudine M.,  
 RT "Conservation of sequence and structure flanking the mouse and human  
 beta-globin loci: the beta-globin genes are embedded within an array  
 of odorant receptor genes."  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:5129-5134(1999).  
 RN [2]  
 RP ERRATUM.  
 RA Bulger M., van Doorninck J.H., Saitoh N., Telling A., Farrell C.,  
 RA Bulger M.A., Felsenfeld G., Axel R., Groudine M.,  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:8307-8307(1999).  
 RN [3]  
 RP SEQUENCE OF 66-283 FROM N.A.  
 RA Fuchs T., Malecova B., Linhart C., Sharran R., Khen M., Herwig R.,  
 RA Shmulevich D., Elkon R., Steinfath M., O'Brien J.K., Radelof U.,  
 RA Lehrach H., Olander Z., Glusman G., Lancel D., Shahr R.,  
 RT "DEFOG: a practical scheme for deciphering families of genes."  
 Submitted (JUL-2001) to the EMBL/GenBank/DBS databases.  
 CC -1- FUNCTION: PUTATIVE ODOURANT RECEPTOR.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF137396; AAK29425.2; -  
 DR EMBL: AF137396; AAK29425.2; -  
 DR Gene; HGNC:14703; OR51B2.  
 DR GO: GO:0016021; C: integral to membrane; NAS.  
 DR GO: GO:0004984; F: olfactory receptor activity; NAS.  
 DR GO: GO:0007608; P: olfaction; NAS.  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR Pfam: PF00001; 7tm.1.1  
 DR PRINTS: PR000237; GPCR\_RHODOPSN.  
 DR PROSITE: PS00237; G-PROTEIN RECP\_F1.1; 1.  
 DR PROSITE: PS00262; G-PROTEIN RECP\_F1.2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction.  
 FT DOMAIN 1 23  
 FT TRANSMEM 24 47  
 FT DOMAIN 48 55  
 FT TRANSMEM 56 77  
 FT DOMAIN 78 98  
 FT TRANSMEM 99 118  
 FT DOMAIN 119 137  
 FT TRANSMEM 138 156  
 FT DOMAIN 157 193  
 FT TRANSMEM 194 217  
 FT DOMAIN 218 234  
 FT TRANSMEM 235 257  
 FT DOMAIN 258 272  
 FT TRANSMEM 273 292  
 FT DOMAIN 293 312  
 FT DISULFID 95 187  
 FT CARBOHYD 4 4  
 SQ SEQUENCE 312 AA; 35513 MW; 0DA116026778EDD CRC64;  
 Query Match 45.1%; Score 728.5; DB 1; Length 312;

Best Local Similarity 44.1%; Pred. No. 2,7e-49;  
 Matches 132; Conservative 76; Mismatches 90; Indels 1; Gaps 1;  
 QY 10 SATYFLLGPGLEAOPFLAPLCSLYLAVALGNTLIYVTSIHPRKPIFLCMS 69  
 DB 6 TAAPELTPGPGLEAHHMISIPFEVAVYCTILGNMELLILHDSLHEPMYPTLMA 65  
 QY 70 GIDILISTSPMPKALIFWENSTTIQDACLQIFAIHSIGMESTVLLAMAFRYAIC 129  
 DB 66 GTDLMTVLTMPVWGLWNVNREISVCGFLQAYFIHSLSVSGSLAMAYDRFAIR 125  
 QY 130 HPLRRAVLTLPVYTIQVAAVRGALAPLDPVFLKQLPFCNSNLISHYCHQVMKL 189  
 DB 126 NPLRYASIFNTNTVIALGVGVFLRGVSLIPVLLRFSFSYCKSHVITRAFCHQEMRL 185  
 QY 190 ACDIDRVNVAVGLIIVISAIGDLSLISFSLILKTVLGL-TREDAQKAFGTVSHVCA 248  
 DB 186 ACADITFNRILPYVILSLITFLDLSLILSYILNLTVIGAGGERAKALNTCISHSC 245  
 QY 249 VLFYVYVPIGLSVNHRSPKRDSPPLVILANLYLVPLVPLNPIYGVKTKERQRLRL 307  
 DB 246 VLFYVYVWGLTPIYRFGKNVPEVHIMSYIFELPEPLMNPVYSIKTKQIQYGIIRL 304  
 RESULT 10  
 OYB6 HUMAN STANDARD; PRT; 313 AA.  
 AC Q96RD3;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Olfactory receptor 52B6.  
 GN OR52B6.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 70-287 FROM N.A.  
 RA Fuchs T., Malecova B., Linhart C., Sharran R., Khen M., Herwig R.,  
 RA Shmulevich D., Elkon R., Steinfath M., O'Brien J.K., Radelof U.,  
 RA Lehrach H., Olander Z., Glusman G., Lancel D., Shahr R.,  
 RT "DEFOG: a practical scheme for deciphering families of genes."  
 Submitted (JUL-2001) to the EMBL/GenBank/DBS databases.  
 CC -1- FUNCTION: PUTATIVE ODOURANT RECEPTOR.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF137396; AAK29425.2; -  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR Pfam: PF00001; 7tm.1.1  
 DR PROSITE: PS00237; G-PROTEIN RECP\_F1.1; 1.  
 DR PROSITE: PS00262; G-PROTEIN RECP\_F1.2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction.  
 FT DOMAIN 1 27  
 FT TRANSMEM 28 51  
 FT DOMAIN 52 59  
 FT TRANSMEM 60 81  
 FT DOMAIN 82 102  
 FT TRANSMEM 103 122  
 FT DOMAIN 123 144  
 FT TRANSMEM 142 160  
 FT DOMAIN 161 197  
 FT TRANSMEM 198 220

FT	DOMAIN	221	237	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	238	260	6 (POTENTIAL).
FT	DOMAIN	261	275	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	276	295	7 (POTENTIAL).
FT	DOMAIN	296	313	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	5	5	N-LINKED (GLUCNAC. . .) (POTENTIAL).
SO	SEQUENCE	313 AA;	35523 MW;	6FC17ADBC0B20B83 CRC64;
Query Match				
Best Local Similarity		43.5%;	Score 704;	DB 1; Length 313;
Matches 139;		Conservative 57;	Mismatches 112;	Indels 2; Gaps 2
QY	5	NGNESSATYFILGLDGLDEEAOFWLAPFLCSLYLIAVGNLTIIYIVRTSHSLHEPMYIF	64	
DB	5	NDNQFPTSSFLGLDGLDEEAOFWLAPFLCSLYLIAVGNLTIIYIVRTSHSLHEPMYIF	64	
QY	65	LCMLSGDILISTSSPMKMLAFWNSTTIQDACLQIPIAHSLSGMESTVLLMAAPDR	124	
DB	65	LANLDDIDLSLSTATIPKMLGFWNINKEISFGSLSMFPIHFTMEIYLVMAAPDR	124	
QY	125	YVAICHPLEHAYVLIPRVTKIGVAAYVGAALMAPELVPIKQLPGRSNILSHSYCLHQ	184	
DB	125	YVACICPPLWYMLTSLKISLISLAGAVLRSLYMYPLVFLRLPFCGHRILIPHYCEHM	184	
QY	185	DVAKLACDDIRVVVVVGLIVIIISAIGLDSLISPSYLLIKTVYGL-TRBAQKARQTCV	243	
DB	185	GIRALKACASIKRINIMG-LGSISLLDLVLLISHILIIYAVCLPSWAKRIKALANTGC	243	
QY	244	SHVCAVFITYVPFPIGLSMVHRSKRRDSPLYLIANTYLLVPVLANIYVGVTKXEIKOR	303	
DB	244	SHGVILASTPAFSPFPTFCGHDIPOYIHIFLANLYVVVPLNPFVIVGVTKXIRET	303	
QY	304	ILRLFFVATH	313	
DB	304	VLRIFPKTDH	313	
RESULT 11				
ID	OXB4_HUMAN	STANDARD;	PRT;	310 AA.
AC	OQY5F0;			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	Olfactory receptor 51B4 (HORS/Beta1).			
GN	OR51B4.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI Taxid=9606;			
RN	[1] _			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99238494; PubMed=10220430;			
RA	Bulger M., van Doorninck J.H., Saitoh N., Telling A., Farrell C.M.,			
RA	Bender M.A., Felsenfeld G., Axel R., Groudine M.;			
RT	"Conservation of sequence and structure flanking the mouse and human			
RT	beta-globin loci: the beta-globin genes are embedded within an array			
RT	of odorant receptor genes.";			
Proc. Natl. Acad. Sci. U.S.A.	96:5129-5134(1999).			
RN	[2]			
RP	ERRATUM.			
RA	Bulger M., van Doorninck J.H., Saitoh N., Telling A., Farrell C.M.,			
RA	Bender M.A., Felsenfeld G., Axel R., Groudine M.;			
Proc. Natl. Acad. Sci. U.S.A.	96:8307-8307(1999).			
RN	[3]			
REV	REVISIONS.			
RX	MEDLINE=20570519; PubMed=11121057;			
RA	Bulger M., Bender M.A., van Doorninck J.H., Wertzman B., Farrell C.M.,			
RA	Felsenfeld G., Groudine M., Hardison R.;			
RT	"Comparative structural and functional analysis of the olfactory			
RT	receptor genes flanking the human and mouse beta-globin gene			
RT	clusters.";			
Proc. Natl. Acad. Sci. U.S.A.	97:14560-14565(2000).			

```

CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF137356; AA029426.2; -
DR Genew; HGNC:14708; OR51B4.
DR GO; GO:0016021; C: integral to membrane; NAS.
DR GO; GO:0004984; F: olfactory receptor activity; NAS.
DR GO; GO:0007608; P: olfaction; NAS.
DR InterPro; IPR002676; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PRO0237; GPCRHOPOSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; FALSE NEG.
DR PROSITE; PS50262; G-PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Olfaction.
FT DOMAIN 1 23
FT TRANSSEM 24 47
FT TRANSSEM 48 55
FT TRANSSEM 56 77
FT TRANSSEM 78 98
FT TRANSSEM 99 117
FT TRANSSEM 118 136
FT TRANSSEM 137 155
FT TRANSSEM 156 192
FT TRANSSEM 193 216
FT TRANSSEM 217 233
FT TRANSSEM 234 256
FT DOMAIN 257 271
FT TRANSSEM 272 291
FT DOMAIN 292 310
FT DISULFID 95 186
FT CARBOHYD 4 4
SQ SEQUENCE 310 AA; 34926 MW; BFE58363BA2EED67 CRC64; (POTENTIAL).

Query Match 39.1%; Score 632; DB 1; Length 310;
Best Local Similarity 40.5%; Pred. No. 7,4e-42;
Matches 122; Conservative 72; Mismatches 105; Indels 2; Gaps 2

9 SSATYFLLIGLPGLEAQAQFWIAFLGSLYLAVLGNLTIIYVTRTEHSHBPMYIFLQML 68
5 NSAGPFIITGLPGLEAQAQFWIAFLGSLYLAVLGNLTIIYVTRTEHSHBPMYIFLQML 64
69 SGIDILSTSMPEMLAIFMNSTTIOFDACLOIFPHISLGSSESTYLLAMPDRAVAI 128
65 ADYDLGNTFTMPFVLGVLIDQGEIHAACFTOSF-IHSLATVSEGLLVLAIDCFNAI 123
129 CHPLRAHATVLTLPVETKIGVAAVAGALMAAPLPEVFIKQDFPCRSNILLSYCLHQDYMK 188
124 RTPLRVNCIILNLSVNNIGLVLMRGMSILPILISLCYCYGCSRALHTFCLHQDYIK 183
189 LACGDIVNVYGGIIVLISALIGDSLILSSYLLIKTVLGL-TRDAQKAFGTCVSHVC 247
184 LACADITFNHTIYPIIQSLVFLDALIIRSYLLIKTVMGASAGQEPKSLNITCVSHIS 243
248 AVETIYVPIGLSNVHRFSKRRDSPVLITANIYLVPEVNLPIYVGKTEIKRQRIIRL 307
244 CVLVEHITVWGLSIFHRFGKAPVAVVETIWSYVHFLPPFPVNPPIYSIKTKQIORSIIRL 303
308 F 308
304 F 304

```

ID	2D	2D2	HUMAN	STANDARD	PRT	319 AA
AC	Q9N0N1					
DT	16-OCT-2001	(Rel. 40, Created)				
DT	16-OCT-2001	(Rel. 40, Last sequence update)				
DT	15-SEP-2003	(Rel. 42, Last annotation update)				
DE	Olfactory receptor 292.					
GN	OR252.					
OS	Homo sapiens (Human).					
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
OK	NCBI_taxid=9606;					
RN	[1]					
RP	SEQUENCE OF 11-319 FROM N.A.					
RA	Cordy N.;					
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.					
CC	- FUNCTION: PUTATIVE ODORANT RECEPTOR.					
CC	- SUBCELLULAR LOCATION: Integral membrane protein.					
CC	- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (see <a href="http://www.isb-sdb.ch/announce">http://www.isb-sdb.ch/announce</a>					
CC	or send an email to <a href="mailto:license@sdb-sdb.ch">license@sdb-sdb.ch</a> ).					
CC	-----					
DR	EMBL; AL135841; CAB96728.1; -					
DR	GeneW; HGNC:8276; OR252.					
DR	GO; GO:0004984; F:olfactory receptor activity; TAS.					
DR	InterPro; IPR000276; GPCR_Rhodopsn.					
DR	Pfam; PF00001; 7tm1.1					
DR	PRINTS; PR00237; GPCR_RHODOPSN.					
DR	PROSITE; PS00237; G-PROTEIN RECP_F1.1; 1.					
DR	PROSITE; PS50262; G-PROTEIN RECP_F1.2; 1.					
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;					
KW	MultiGene family; Olfaction.					
FT	DOMAIN	1	26			
FT	TRANSMEM	27	58			
FT	TRANSMEM	51	58			
FT	TRANSMEM	59	80			
FT	DOMAIN	81	101			
FT	TRANSMEM	102	121			
FT	DOMAIN	122	140			
FT	TRANSMEM	141	159			
FT	DOMAIN	160	196			
FT	TRANSMEM	197	220			
FT	DOMAIN	221	237			
FT	TRANSMEM	238	260			
FT	DOMAIN	261	279			
FT	TRANSMEM	280	299			
FT	DOMAIN	300	319			
FT	DISULFID	98	190			
FT	CARBOHYD	5	5			
SQ	SEQUENCE	319 AA;	35172 MW;	602602FCA96A179F	CRC64;	
Query Match		28.7%;	Score 464;	DB 1;	Length 319;	
Best Local Similarity		35.3%;	Pred. No. 6.6e-29;			
Matches 113;	Conservative	64;	Mismatches 11;	Indels	32;	Gaps 9;
OY	7	NESSATY-FILHGL--PGLBAQWALFPCSLYLIVLGNLTIIVVTRHSHHEMY	62			
DB	5	NETSYVMFVLLRLSAHPELEKTFVL--ILMLVILLNGVILLVTIDSRILHTMY	61			
OY	63	IFLCNLGIDILISTSSMPKMLAFWENSTTIQPDALLQIFAHISLGSESTVLLAAAF	122			
DB	62	FFLGMLSLDIDCIPTSSVPLVDSFLTRQETISFSCACVMAALSRAMAGTCLILSMNAF	121			
OY	123	DRVYALGRLHAATVLLPRYTKIGVA--VVRGALMLAPLPFTKQLPFCRSNIIISHSYC	181			
DB	122	DRVYALGRLKYSYVMKAAVMPMAASSWALIGASVYHTSLATQLPFCGDNVYHETC	180			

[illegible]



SQ SEQUENCE 312 AA; 34333 MW; 8D3877EBB51E132 CRC64;  
 Query Match 27.1%; Score 439; DB 1; Length 312;  
 Best Local Similarity 34.2%; Pred. No. 5.5e-27;  
 Matches 111; Conservative 67; Mismatches 123; Indels 24; Gaps 11;  
 CC 2 VDPNGNSSATYFLLIGL---PGLLEAOFWIAFPLCSLYLAVALGNLTIIYVREHSH 58  
 DB 1 MEVSNSSSGT-FILMGVSDHPHLEIFFAV---ILASVLLTVGNLTIIILSRDLARH 56  
 CC 59 EPMYIFLCMSGIDILLSTSSMPKMLAIFWNSTTIOFDCLLOIPIAHSLSMESTVLL 118  
 DB 57 TPTMFFLSNLSLDAITSSVPMKMLWGPDKTISYGGCVQLYVFLWLGATECILLV 116  
 CC 119 AMAFDRYVALCHPRHATVLTLPRTKIGVAAV--VRGALMALPLVFIKOLPFCRSNLI 176  
 DB 117 VMADRYVAVACRPLHYMTVMN-PLIC-WGLALISWGLGNSVIGSTFTLLQLPFCGRHY 174  
 CC 177 SHSICLHODVWKLACDDIRVN--VYGLIVITSAIGDSLISFSLIKTVLGL-TRR 233  
 DB 175 DNFLCEVPAMIKLACGDTSENALVNGVCTFFVVPVSVLV--SYCFIAQAVMKIRSYE 232  
 CC 234 AQAFAFCVSHVCVAFIYVPIGLSMVHRF--SKRSDPLPIVILANITLVPEVLP 290  
 DB 233 GRRAFPTCVSHLVVFLFYG-----SAICYLLPAKSSNQSQKTSISLFSVITPMNP 287  
 CC 291 IVGVKTEIRQIRLRLPHVATHAS 315  
 DB 288 LITLRNKEVKGALGRLLGKRGAS 312  
 RESULT 14  
 OLFA\_RAT STANDARD; PRT; 312 AA.  
 ID OLFA\_RAT STANDARD; PRT; 312 AA.  
 AC P23273;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 01-JUL-1993 (Rel. 26, Last annotation update)  
 DE Olfactory receptor-like protein I14.  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CC NCBI\_TaxID=10116;  
 CC [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91191556; PubMed=1840504;  
 RA Buck L., Axel R.;  
 RT "A novel multigene family may encode odorant receptors: a molecular  
 RT basis for odor recognition.";  
 RL Cell 65:175-187 (1991).  
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: OLFACTORY EPITHELIUM.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; M64391; AAA41754.1; -  
 CC PIR; I23701; I23701.  
 CC InterPro; IPR00276; GPCR\_Rhodopsin.  
 CC Pfam; PF00001; 7tm.1; 1.  
 CC PRINTS; PR00237; GPCRHOPOPSN.  
 CC PROSITE; PS00237; G-PROTEIN\_RECP\_P1.1; 1.  
 CC PROSITE; PS50262; G-PROTEIN\_RECP\_P1.2; 1.  
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
 CC Multigene family; Olfaction.  
 CC DOMAIN 1 25 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 26 49  
 FT DOMAIN 50 57  
 FT TRANSMEM 58 79  
 FT TRANSMEM 80 100  
 FT TRANSMEM 101 120  
 FT TRANSMEM 121 139  
 FT TRANSMEM 140 158  
 FT TRANSMEM 159 196  
 FT TRANSMEM 197 219  
 FT TRANSMEM 220 236  
 FT TRANSMEM 237 260  
 FT TRANSMEM 261 272  
 FT TRANSMEM 273 292  
 FT TRANSMEM 293 312  
 FT CARBOHYD 5 5  
 FT DISULFID 97 189  
 SQ SEQUENCE 312 AA; 35718 MW; 08F66546994B1680 CRC64;  
 Query Match 26.9%; Score 434.5; DB 1; Length 312;  
 Best Local Similarity 33.5%; Pred. No. 1.2e-26;  
 Matches 105; Conservative 64; Mismatches 119; Indels 25; Gaps 7;  
 CC 7 NSSATYFLLIGLGLLEAOFWIAFPLCSLYLAVALGNLTIIYVREHSHHEMYTFLC 66  
 DB 5 NQTLIEFLGLDLPISRYHLLFPLALPLAKLTIIIGNLITIVLRDSHLHMPYIFLS 64  
 CC 67 MLSGIDILISTSSMPKMLAIFWNSTTIOFDCLLOIPIAHSLSMESTVLLAAFDPRV 126  
 DB 65 NLSFSDLCFSSVTPMKLLQNMQSQVPSISYGLCTQLYFVWFEDMESFLIVVAYRYV 124  
 CC 127 AICPHRAVATLTLPRTKIGVAAV-----RAALMAPLPVRIKOLPFCRSNLSHY 180  
 DB 125 AICPEPLRTITMS---TRKASLVLTLMLMTHTALHLTL--LIARSLFCKNVIIHFF 178  
 CC 181 CHQDVWKLACDDIRVN--VYGLIVITSAIGDSLISFSLIKTVLGL-TRR 234  
 DB 179 CDISALKLCSQSDIYVNEMLYIIGLIIII-----PFLIVMSYVAFISILKFPISIQD 233  
 CC 235 QAKAFGCVSHVCVAFIYVPIGLSMVHRF--SKRSDPLPIVILANITLVPEVLPVYV 294  
 DB 234 IYKVFSTGSHLSVLTLPFYGTISYVLC--SGNNTVKEIAMAMVTVVTPMLNPITS 291  
 CC 295 VKTEIRQIRLRL 307  
 DB 292 LNRDMKRALIRV 304  
 RESULT 15  
 O2C1\_HUMAN STANDARD; PRT; 312 AA.  
 ID O2C1\_HUMAN STANDARD; PRT; 312 AA.  
 AC O95371;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Olfactory receptor 2C1 (Olfm13).  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 CC NCBI\_TaxID=9606;  
 CC [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99063798; PubMed=9847080;  
 RA Centola M., Chen X., Sood R., Deng Z., Aksevitjevich I., Blake T.,  
 RA Riecke D., Chen X., Wood G., Zaks N., Richards N., Krizman D.,  
 RA Mansfield E., Apostolou S., Liu J., Shafran N., Vedula A., Hamon M.,  
 RA Cerecek A., Kahan T., Gumucio D., Callen D.F., Richards R.I.,  
 RA Moyzis R.K., Doggett N.A., Collins F.S., Liu P.P.,  
 RA Fischer-Grodjan N., Kastner D.L., 700 kb transcript map around the  
 RA "Construction of an approximately 700 kb transcript map around the  
 RA familial Mediterranean fever locus on human chromosome 16p13.3.";  
 RL Genome Res. 9:1-21 (1998).  
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.

```
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF098664; AAC83557.1; -.
CC Genew: HGNC:8242; OR2C1.
CC GO: GO:0004984; P:olfactory receptor activity; TAS.
CC InterPro: IPR000276; GPCR_Rhodopsn.
CC Pfam: PF00001; 7tm.1; 1.
CC PRINTS: PR00237; GPCR_RHODOPSIN.
CC PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE: PS00282; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Multigene family; Olfaction.
CC FT DOMAIN 1 25
CC FT TRANSMEM 25 49 1 EXTRACELLULAR (POTENTIAL).
CC FT DOMAIN 50 57 2 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 58 79 2 (POTENTIAL).
CC FT DOMAIN 80 100 3 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 101 120 3 (POTENTIAL).
CC FT DOMAIN 121 139 4 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 140 158 4 (POTENTIAL).
CC FT DOMAIN 159 196 5 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 197 219 5 (POTENTIAL).
CC FT DOMAIN 220 236 6 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 237 259 6 (POTENTIAL).
CC FT DOMAIN 260 272 7 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 273 292 7 (POTENTIAL).
CC FT DOMAIN 293 312 7 CYTOPLASMIC (POTENTIAL).
CC FT DISULFID 97 189 BY SIMILARITY.
CC FT CARBOHYD 5 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 312 AA; 34506 MW; CC4D7D349BB8B9E7 CRC64;

Query Match 26.7%; Score 432.5; DB 1; Length 312;
Best Local Similarity 33.4%; Pred. No. 1.7e-26;
Matches 102; Conservative 58; Mismatches 136; Indels 9; Gaps 5;

QY 7 NESAATYFLILCPGLEERQFWLAPLGLSVLIATVLSGNTIIVYRTSHLHEPNTIFLC 66
DB 5 NDSSTLQGFVLMKSISSHPOLEMTFFAILFSYLLTLGNSITILLRLPARLHTPMYFFLS 64
QY 67 MLSGIDILISTSMPEKMLAIFWENSTTIOFDACLIQIFAISLSGMESTVLLMAFDRYV 126
DB 65 NLSLIDLAFAITSSVQMLINLMGPKITISYGGCIQQLVFLMGLATECILLVMAFDRYV 124
QY 127 AICHPLEAATVTLTPRVTKIGVAAVVRGAALMAPLPVFIKQIPGCRNLSHSYCLHODV 186
DB 125 AAVCPLEAATVTLTPRVTKIGVAAVVRGAALMAPLPVFIKQIPGCRNLSHSYCLHODV 184
QY 187 MSLACDPIRVN--VYVGLIVITSAIGLDSLISFSYLLITKTVLGL--TREAOAKAFGRVY 243
DB 185 IRLAGDITSLNQAVALNGVCTFTAVPLSIIV--SYCLIAQAVLKIHSAEGRRKAFNTCL 242
QY 244 SHVCAVFITY-VFTGLSNVHRFSRRDSPLEVIITANIYLLVPPVLANIVYGVKTKEIRQ 302
DB 243 SHLVVFFLFYGSASVGYLLPAKNSKODGK--FSLFYSIVTPTMVPNPLIYTLNMEVKG 299
QY 303 RILRL 307
DB 300 ALRRL 304
```

Search completed: February 9, 2004, 16:24:16  
Job time : 18 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 16:22:06 / Search time 20 Seconds  
(without alignments)  
1524.274 Million cell updates/sec

Title: US-10-017-066a-2

Sequence: 1 MVDPMNGESSATYFLILGHP.....KEIRQRILRLPVAATHASEP 317

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

1: PIR 76: \*  
2: PIR1: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	552	34.1	292	2	A60547
2	473.5	29.3	313	2	E45774
3	453	28.0	307	2	S29710
4	439	27.1	312	2	A46247
5	434.5	26.9	312	2	I23701
6	431.5	26.7	310	2	E23701
7	430.5	26.6	319	2	UC5624
8	429.5	26.6	312	2	S29708
9	428	26.5	328	2	G45774
10	427.5	26.4	314	2	S20572
11	423.5	26.3	314	2	A37286
12	423.5	26.2	314	2	S29707
13	420.5	26.0	321	2	H45774
14	418.5	25.9	305	2	S29711
15	418.5	25.9	309	1	SS1356
16	413	25.5	312	2	A48413
17	411	25.4	160	2	SS8066
18	411	25.4	344	2	I45774
19	407.5	25.2	311	2	UC5200
20	407	25.2	313	2	S20571
21	404.5	25.0	312	2	G23701
22	404.5	25.0	313	2	B23701
23	402.5	24.9	311	2	C23701
24	402.5	24.9	314	2	H23701
25	401	24.8	317	2	A45774
26	396.5	24.5	316	2	A57069
27	395.5	24.5	328	2	A45774
28	395.5	24.5	328	2	D45774
29	393.5	24.3	328	2	B45774

#### ALIGNMENTS

30	390.5	24.1	320	2	S20573	olfactory receptor
31	388.5	24.0	312	2	A46750	olfactory receptor
32	386.5	23.9	304	2	S29709	olfactory receptor
33	384.5	23.8	315	2	UC5836	olfactory receptor
34	383.5	23.7	328	2	C45774	olfactory receptor
35	383.5	23.7	333	2	A23701	olfactory receptor
36	379.5	23.5	327	2	F23701	olfactory receptor
37	375	23.2	317	2	D23701	olfactory receptor
38	370	22.9	315	2	UC5201	chemoreceptor TB56
39	368.5	22.8	315	2	UC4658	olfactory receptor
40	366	22.6	318	2	UC5202	chemoreceptor TB64
41	313.5	19.4	234	2	S29000	G protein-coupled
42	310.5	19.2	222	2	B40745	odorant receptor
43	303.5	18.8	264	2	PC4369	olfactory receptor
44	289.5	17.9	234	2	S28999	G protein-coupled
45	285.5	17.7	222	2	D40745	odorant receptor

#### RESULT 1

A60547  
hypochemical protein (HPRH breakpoint 3' region) - human (fragment)  
N/Alternate names: olfactory receptor homolog  
C/Species: Homo sapiens (man)  
C/Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #ext\_change 26-Aug-1999  
A/Accession: A60547; A60482  
R/Feingold, E.A.; Forget, B.G.  
Blood 74, 2178-2186, 1989  
A/Title: The breakpoint of a large deletion causing hereditary persistence of fetal hemoglobin  
A/Reference number: A60547; PMID:90028751; PMID:2478223  
A/Accession: A60547  
A/Molecule type: DNA  
A/Residues: 1-292 <FEI>  
C/Superfamily: olfactory receptor OR14

Query Match 34.1%; Score 552; DB 2; Length 292;

Best Local Similarity 40.3%; Pred. No. 2.2e-41;

Matches 114; Conservative 49; Mismatches 88; Indels 32; Gaps 6;

Qy	16	LIGHPLGEAQAQFWLAPLGLVLAVALGNTIIVITRSHLHPMYFLMGIDIL 75	
Db	29	LVGPGLESVQCWIGPFCALYILAMINGNLSIKSRSLHPIYIFKMGARDIAL 88	
Qy	76	STSGMPKMLAIFPNSTTIOFDACLIQIF-----AIH-SLSGMESTVLLAMFDPYVAI 128	
Db	89	ASSIMPMGLGSGU-----MCLKSILIPACENVVHPHIAIGESGLVAMALDRYVAI 140	
Qy	129	CHPLRHATVTLPRVTIKGVAAYVGAALMAAPVPIK-OLPFCRANILSHSYCLHQDVN 187	
Db	141	CYPLRHANITFHQLVIGTMVTLRAALIVAPCHVLIKRFQYHTTVISHSCENMAIV 200	
Qy	188	KLACDDIRNVVVGGLVYIISALIGDSLISFSYLLIKTYLGL-TEDQAKAGTCVSHV 246	
Db	201	KLAANYQVAKIKGLFAFVAGFDLFTLLSYIQIFIVFRLPQEARPKAFTICIAH 260	
Qy	247	CAVIFVYVPFPIGLSMVRFKSRKDSPLVILANIVLVPPVLN 289	
Db	261	CVLQFYLLAFPSFFTHRFSSH-----IPPIYH 288	

#### RESULT 2

E45774  
odorant receptor 8 - channel catfish  
C/Species: Ictalurus punctatus (channel catfish)  
C/Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #ext\_change 26-Aug-1999  
A/Accession: E45774  
R/Neel, J.; Dowling, M.M.; Buck, L.; Axel, R.; Chess, A.  
Cell 72, 657-666, 1993  
A/Title: The family of genes encoding odorant receptors in the channel catfish.  
A/Reference number: A45774; PMID:93201590; PMID:7916654  
A/Accession: E45774

A:Status: preliminary; not compared with conceptual translation  
A:Structure type: mRNA  
A:Molecule type: mRNA  
A:Residues: 1-313 <N>A>  
A:Experimental source: olfactory epithelium  
A:Note: sequence extracted from NCBI backbone (NCBI:P127748)  
C:Superfamily: Olfactory receptor OR14  
C:Keywords: Olfaction; transmembrane protein

Query Match	29.3%;	Score 473.5;	DB 2;	Length 313;
Best Local Similarity	34.1%;	Pred. No. 2.1e-34;		
Matches 105;	Conservative 66;	Mismatches 132;	Indels 5;	Gaps 4

QY	1	MVDNGNESSATYFIILIGLPGLEEAQPMFLAFLCSLYLLAVGNTITTYVYRTSHSLHEP	60
Db	1	MLAVQNIISFTT-FTLIGFHDLSGKPPILISFLMFLPILSTSNNTLLYLLIISPRALHSP	59
QY	61	MYFLCNLSGIDILLISTSSMPKXLAIFWNSSTTIOFDACLOIFAHSLSGMSSTVLLAM	120
Db	60	MCILIGLMAVVDLSMPFICVPPNMLLSFLFNWKGISIVSGCLVQMFCIHQAGTFQSTILLMM	119
QY	121	AFDEYVAICGFLPHATVLTLPRTKISGVAAVARGAALMAPVDFEIKOLPQRSNILLSHY	180
Db	120	ALDFEFAICGFLPYQKMGKMPNFKITIFPVIRNLEFITTIVSMAGKLTPEINIBDHCV	179
QY	181	CLHODVKKLACDDIRANVAVGLIVITISAIGDLSLISPSYLLIKTVLTGLTREAOARFG	240
Db	180	CEHNAITQIALCGDISINMLIGLTVLDTTAAPIFITTSYIVILVSIL-REGSKACLRVN	238
QY	241	TCVSHVCAVPIFYVVPFGLSMWRFSKRDSP-L-VITANITVLPVPLNANIVGVTKK	299
Db	239	TCITHIVMTVSLTFPLNIAFLSYRI-RNFSRSSKVFSLTMVLPFSPCNIIIVGATKE	296
QY	300	IRQRIQLRL	307
Db	297	IRQQLFLTL	304

```

RESULT 3
S29710
c|factory receptor OR18 - rat
c|Species: Rattus norvegicus (Norway rat)
c|Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 04-Sep-1998
c|Accession: S29710
R:Rattus, K.; Krieger, J.; Strechmann, J.; Boekhoff, I.; Kubick, S.; Baumstark, C.; Breer
A:Title: Cloning and expression of odorant receptors.
A:Reference number: S29707; MUID:93149273; PMID:7678922
A:Accession: S29710
A:Molecule type: mRNA
A:Residues: 1-507 <RAN>
A:Superfamily: olfactory receptor OR14

```

Query Match	28.0%;	Score 453;	DB 2;	Length 307;
Best Local Similarity	33.6%;	Pred. No. 1.4e-32;		
Matches 107;	Conservative 64;	Mismatches 125;	Indels 22;	Gaps 7;

QY 6 GNESAATYELLIGL---PGLSEAQEWLAFPLCSLYLNAVGNLITTYIRREHSLHEMY 62  
 Db 2 GENNNITTEIFLIGLQDPDGRKALFVLF---LIIYVMMNNLIIIVTVIASPSELGSPMY 58  
 QY 63 IFLOMGSDIDILISTSSWFKMLAIFWENSTTICQFDACILQIFAIHSLSGMESTYLMAAF 122  
 Db 59 FFLASLSLIDLFTSAISPKLIADLLVDQKITSIRACMSQLFIHLRGVDIVILVMAY 118  
 QY 123 DRYAIAICHPLRHAIYVLTIPRYTKIGVAAVVRGALMALPVPFIQOLPPCRSNITISHCYL 182  
 Db 119 DRYAIAICPPLHYLIAMNRVCITILLIPAMTGGFTHSLIQIFYVYNLPFCGENVVIDHFICD 178  
 QY 183 HDVYKLIACDRIKRVNVVYGLIIV---ISAIGDLSLISFSLYTLIKVLTGLTREAOKA 238  
 Db 179 MSPLLVLAICD---TYPIGLTVIANGGVNCLVIFITLLIG-SYGIITSLSKTQSGEGRKA 234  
 QY 239 FGTCVSHVCAVFLIYVP-FIGLSMVRHFSKRRDSPVILIANITYLVLPVPLNIVYGVK 296

Dd	235	LSTGSSHLVILFEPCEIMYAPVNF-----PIDCITVEYTIITPMNLIVTLR	288
Oy	297	TKEIRQILRLFHVATHA	314
Dd	289	NSEIKSCMKLWCKMLHA	306

```

RESULT 4
A#6247
C:factory receptor OR3 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C:Accession: A46247
R:Nef, P.; Hermans-Borgmeyer, I.; Artzt-John, H.; Beasley, L.; D'Amico, V.E.; Heinemann
Proc. Natl. Acad. Sci. U.S.A. 89, 8948-8952, 1992
A>Title: Spatial pattern of receptor expression in the olfactory epithelium.
A:Reference number: A46247; MUID:93028384; PMID:1384038
A:Accession: A46247
A>Status: Preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-312 <NEF>
A:Cross-references: GB:MM4005; NID:G200153; PIND:AAA9862.1; PID:G200154
A>Note: sequence extracted from NCBI backbone (NCBIP:115362)
C:Superfamily: olfactory receptor OR14

```

Query Match	27.1%;	Score 439;	DB 2;	Length 312;
Best Local Similarity	34.2%;	Pred. No. 2.4e-31;		
Matches 111;	Conservative 67;	Mismatches 123;	Indels 24;	Gaps 11

[illegible]

```

RESULT 5
123701
Olfactory receptor 114 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 26-Aug-1999
C:Accession: 123701
R:Buck, L.; Axel, R.
Cell 65, 175-187, 1991
A:Title: A novel multigene family may encode odorant receptors: a molecular basis for od
A:Reference numbers: A23701; MUID:51151556; PMID:1840504
A:Accession: 123701
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-312 <BNC>
A:Cross-references: GB:W64391; NID:G205843; PIDN:AAA41754.1; PID:G205844
C:Superfamily: Olfactory receptor OR14
Keywords: G protein-coupled receptor; transmembrane protein

```

Query Match 26.9%; Score 434.5; DB 2; Length 312;  
 Best Local Similarity 33.5%; Pred. No. 6e-31;  
 Matches 105; Conservative 64; Mismatches 119; Indels 25; Gaps 7;

QY 7 NESAATYFILGLPGLBEAOFWLAFLPCLSYLIANLVGNLTIIYVRETHSLHEPMYIFLC 66  
 Db 5 NQTLLEFLLLGLPPESEYHLLFYALFLAYLTIILNLLIIVLRDLSHLMHMYIFLS 64  
 QY 67 MLGSDILISTSSMPKMLAIFWNSTTIQFACLLQIFAHSLSGMESTVLLAMAFRY 126  
 Db 65 NLSPDLCPSSVTMFKLQNMOSQPSISYGCITQIFKFWEDMSFLVMAIDRY 124  
 QY 127 AICHELRAATVLTLPRTKIGVAV-----RGAALMAPLPVEIKQIPCRSNILSHSY 180  
 Db 125 AICPELRATYIMS-----TKFCASVLLMLMTMHLLHTL--LIARLSFCEKQVILHFF 178  
 QY 181 CLHQVMKLACDDIRVN-----VYGLIVIIISAIGDLSLISFSYLLIKTVGL--TREA 234  
 Db 179 CDISALLKLSGSDIYVNLMTYIILGGLIITL-----PFLIIVSYVRIFFSILKFPSTIQD 233  
 QY 235 QAKAVGTCVSHVCAVFIYFVPIGLSMVHRFSKRDSPLPVIILANIYLVPPVLPVYV 234  
 Db 234 IYKAVSTGSHLSVLTLEFGTIFGILCP--SGNNSYVKEILAMAMMTVTVPMLNPTYS 291  
 QY 295 VKTKRIQRIRL 307  
 Db 292 LNRDMKRALIRV 304

## RESULT 6

E23701  
 Olfactory receptor 13 - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 26-Aug-1999  
 C/Accession: E23701  
 R/Buck, L.; Axel, R.  
 Cell 65, 175-187, 1991  
 A/Title: A novel multigene family may encode odorant receptors: a molecular basis for od  
 A/Reference number: A23701; MUID:91191556; PMID:1840504  
 A/Accession: E23701  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-310 <BUC>  
 A/Cross-references: GB:M64385; NID:g205831; PIDN:AAA41748.1; PID:g205832  
 C/Superfamily: olfactory receptor OR14  
 C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 26.7%; Score 431.5; DB 2; Length 310;  
 Best Local Similarity 33.2%; Pred. No. 1.1e-30;  
 Matches 102; Conservative 71; Mismatches 12; Indels 13; Gaps 7;

QY 7 NESAATYFILGLPGLBEAOFWLAFLPCLSYLIANLVGNLTIIYVRETHSLHEPMYIF 64  
 Db 3 NQTLLEFLLLGLPPESEYHLLFYALFLAYLTIILNLLIIVLRDLSHLMHMYIFLC 60  
 QY 65 LCMISGIDILISTSSMPKMLAIFWNSTTIQFACLLQIFAHSLSGMESTVLLAMAFDR 124  
 Db 61 LSNLSFSDLCFSSVMPKLOMRSDTISIPYGCIAQTFEWMVGDMSFLVMAAYDR 120  
 QY 125 YVAICHPRLHATVLTLPRTKIGVA--AVRGAALMAPLPVEIKQIPCRSNILSHSY 181  
 Db 121 YVAICHPRLHATVLTLPRTKIGVA--AVRGAALMAPLPVEIKQIPCRSNILSHSY 177  
 QY 182 LHODVWLKACDDIRVN-----VYGLIVIIISAIGDLSLISFSYLLIKTVGL--TREA 240  
 Db 178 DLFLVLLKACSDTYINELMIFIMSTLIIIPFLIVMSYARISILKVPSTOGICKVFS 237  
 QY 241 TCVSHVCAVFIYFVPIGLSMVHRFSKRDSPLPVIILANIYLVPPVLPVYVGVTKEI 300  
 Db 238 TCGSHLSVSLFYGILIGVILCP--AGNNSYVKEIVAMMTVTVPMLNPTYSILRNDM 295  
 QY 301 RORIRL 307  
 Db 301 RORIRL 307

Db 296 KRALIRV 302

## RESULT 7

UCS624  
 Olfactory receptor HTPCR92 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 14-Oct-1997 #sequence\_revision 07-Nov-1997 #text\_change 20-Jun-2000  
 C/Accession: UCS624; S58003  
 R/Vanderhaeghen, P.; Schumann, S.; Vassart, G.; Parmentier, M.  
 Biochem. Biophys. Res. Commun. 237, 283-287, 1997  
 A/Title: Molecular cloning and chromosomal mapping of olfactory receptor genes expressed  
 A/Reference number: UCS624; MUID:97415789; PMID:9268701  
 A/Contents: Testis  
 A/Accession: UCS624  
 A/Molecule type: mRNA  
 A/Residues: 1-319 <VAN>  
 A/Cross-references: GB:Y10530; NID:g2792017; PIDN:CAA71558.1; PID:g2792018  
 R/Vanderhaeghen, P.; Schumann, S.; Vassart, G.; Parmentier, M.  
 Submitted to the EMBL Data Library, July 1995  
 A/Description: Male germ cells from several mammalian species express a specific repertoire  
 A/Reference number: S57995  
 A/Accession: S58003  
 A/Molecule type: mRNA  
 A/Residues: 126-282 <VAN>  
 A/Cross-references: EMBL:X89677; NID:g902337; PIDN:CAA61824.1; PID:g902338  
 C/Genetics:  
 A/Map position: 19p13.1  
 C/Function:  
 A/Description: Involved in control of sperm physiology  
 C/Superfamily: olfactory receptor OR14  
 C/Keywords: G protein-coupled receptor; transmembrane protein  
 F/26-46/Domain: transmembrane #status predicted <TM1>  
 F/58-79/Domain: transmembrane #status predicted <TM2>  
 F/101-120/Domain: transmembrane #status predicted <TM3>  
 F/140-164/Domain: transmembrane #status predicted <TM4>  
 F/197-218/Domain: transmembrane #status predicted <TM5>  
 F/237-260/Domain: transmembrane #status predicted <TM6>  
 F/272-292/Domain: transmembrane #status predicted <TM7>

Query Match 26.6%; Score 430.5; DB 2; Length 319;  
 Best Local Similarity 33.3%; Pred. No. 1.4e-30;  
 Matches 102; Conservative 71; Mismatches 118; Indels 15; Gaps 8;

QY 6 GNESAATYFILGLPGLBEAOFWLAFLPCLSYLIANLVGNLTIIYVRETHSLHEPMYIF 62  
 Db 4 GNDTQISEFLILGFSQEPGLQFLGL--FLSMIVLVGNLIIILATISDSHLETPMY 60  
 QY 63 IFCLMLSGIDILISTSSMPKMLAIFWNSTTIQFACLLQIFAHSLSGMESTVLLAMAF 122  
 Db 61 FFLSNLSFADI CVTSTTIPKMLNIOQKRVITVIAICMQMFFILFAQFENFLISVMAY 120  
 QY 123 DRVVAICHELRAATVLTLPRTKIGVAAVRGAALMAPLPVFIKQIPCRSNILSHSY 181  
 Db 121 DRVVAICHELRAATVLTLPRTKIGVAAVRGAALMAPLPVFIKQIPCRSNILSHSY 179  
 QY 182 LHODVWLKACDDIRVN-----VYGLIVIIISAIGDLSLISFSYLLIKTVGL--TREA 238  
 Db 180 ELNOVITGLACSDFLHNMVITFVALIGGPTGIL--YSYKILISSIHAISSAOGKYXA 237  
 QY 239 FGCVSHVCAVFIYFVPIGLSMVHRFSKRDSPLPVIILANIYLVPPVLPVYVGVTK 298  
 Db 238 FSTCASHLSVSLFYGILIGVILCP--SSAATRSHSSATASWMTVTVPMLNPTYSILR 295  
 QY 299 EIRORI 304  
 Db 296 DIRAL 301

## RESULT 8

S29708  
 Olfactory receptor OR12 - rat  
 C/Species: Rattus norvegicus (Norway rat)

C/Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 04-Sep-1998  
 C/Accession: S29708  
 R/Ranking: K.; Krieger, J.; Strotmann, J.; Boehnke, I.; Kubic, S.; Baumstark, C.; Breez  
 Nature 361, 353-356, 1993  
 A/Title: Cloning and expression of odorant receptors.  
 A/Reference number: S29707; MUID:93149273; PMID:7678922  
 A/Accession: S29708  
 A/Molecule type: mRNA  
 A/Residues: 1-312 <RAM>  
 C/Superfamily: olfactory receptor OR14

Query Match 26.6%; Score 429.5; DB 2; Length 312;  
 Best Local Similarity 31.7%; Pred. No. 1.7e-30;  
 Matches 98; Conservative 75; Mismatches 119; Indels 17; Gaps 6;

QY 7 NEESSATYFLIGPGLEAQLAF--PLCSLYLAVGNLTIIYVTEHSHPEMYIF 64  
 DB 5 NQIVISQFLILGP--IPPEHWHLYTLLNLTIIILNLTIIILDSLHIMYIF 62  
 QY 65 LCMISGIDILISTSSMPKMLAFWNSITIQDACLQIFAIHSLSGNESTVLLAMAFDR 124  
 DB 63 LKSLSPDSCFSVTPKILQNMNQDSTITYTGCITQWFSMPAGMEIFLVSMAYDR 122  
 QY 125 YVAICPLHATVLTLPRTKIGVAAYVGAALMAPLVPFIKQLPFGCSNLSHSYCLHQ 184  
 DB 123 YVAICPLHATVLTLSMSPKFCVCSLSWPNVLYSMLHTLLARLSFCNDVIPEFCDIS 182  
 QY 185 DWKLACDDIRVN----VYGLIVISAIGLSLISFSYLLIKTVLG--TREAOAKA 238  
 DB 183 ALTKACSDTYINELMIFLIGLTVI----PFLIIWTVQVICSILTKPSRAIKI 237  
 QY 239 FGTGVSHVCAVFIFFYPIFGLSMWRHRSKRDSPLPILANITYLLVPVLANPIYGVKTK 298  
 DB 238 FSTCGSHLSTVSLFYGTVIGLYLC--SSANSTVKEITWAMNITVTPMLNPFYISLNR 295  
 QY 299 EIRQIRL 307  
 DB 296 DIKELAVR 304

RESULT 9  
 G45774  
 odorant receptor 202 - channel catfish  
 C/Species: Ictalurus punctatus (channel catfish)  
 C/Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 04-Sep-1998  
 C/Accession: G45774  
 R/Rgal, U.; Dowling, M.M.; Buck, L.; Axel, R.; Chess, A.  
 Cell 72, 657-666, 1993  
 A/Title: The family of genes encoding odorant receptors in the channel catfish.  
 A/Reference number: A45774; MUID:93201590; PMID:7916654  
 A/Accession: G45774  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-328 <NGA>  
 A/Experimental source: olfactory epithelium  
 A/Note: sequence extracted from NCBI backbone (NCBIP:127746)  
 C/Superfamily: olfactory receptor OR14  
 C/Keywords: olfaction; transmembrane protein

Query Match 26.5%; Score 428; DB 2; Length 328;  
 Best Local Similarity 29.4%; Pred. No. 2.4e-30;  
 Matches 92; Conservative 76; Mismatches 127; Indels 18; Gaps 7;

QY 4 PNGNESSATYFLIGPGLEAQLAF--APLCSLYLAVGNLTIIYVTEHSHPEMY 62  
 DB 2 PEGNININXNVFVILGPGLPBNVYGLVSMFVYVCTLIGNCTPFLFLREKSLQKDMY 61  
 QY 63 IFCKMSGIDILISTSSMPKMLAFWNSITIQDACLQIFAIHSLSGNESTVLLAMAF 122  
 DB 62 YVLANIAASVLESTTLTKILARVFGSGISFPGCFIQWGFVHPFATNALVAVAF 121  
 QY 123 DRYVAICPLHATVLTLPRTKIGVAAYVGAALMAPLVPFIKQLPFGCSNLSHSYCL 182

DB 122 DRYVAVCNPLRVNIVKESTIILGLCVVSWLBAEPTVLTIVIRATSLPYCASNTVIGCYCD 181  
 QY 183 HQDVKLACDDIRVNVVGLIVITSAIGL--DSLISFSYLLIKTVIGL--TREAOAKA 238  
 DB 182 HAVVTKLACD--RRPYAFPALVSAVWMLPLAFILFSYSIIIVTERTSTGRKTK 238  
 QY 239 FGTGVSHVCAVFIFFYPIFGLSMWRHRSKRDSPLPILANITYLLVPVLANPIY 293  
 DB 239 LSTCSSQLIITITFLPRLCLNYSLSLGIH-----INADIQILVIMLYSLPMPINPVY 293  
 QY 294 GVKTKRQIRL 306  
 DB 294 CLRTKRAKECLKR 306

RESULT 10  
 S20572  
 olfactory receptor - human  
 C/Species: Homo sapiens (man)  
 C/Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 26-Aug-1999  
 C/Accession: S20572  
 R/Parmentier, M.; Libert, F.; Schumann, S.; Schiffrmann, S.; LeFort, A.; Eggerickx, D.;  
 Nature 355, 453-455, 1992  
 A/Title: Expression of members of the putative olfactory receptor gene family in mamma  
 A/Reference number: S20571; MUID:92131132; PMID:1370859  
 A/Accession: S20572  
 A/Status: preliminary  
 A/Molecule type: nucleic acid  
 A/Residues: 1-314 <PAR>  
 A/Cross-references: EMBL:X64994; NID:932085; PID:CAA46127.1; PID:932086  
 C/Superfamily: olfactory receptor OR14  
 C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 26.4%; Score 427.5; DB 2; Length 314;  
 Best Local Similarity 34.5%; Pred. No. 2.5e-30;  
 Matches 108; Conservative 64; Mismatches 120; Indels 21; Gaps 8;

QY 7 NEESSATYFLIGPGLEAQLAF--PLCSLYLAVGNLTIIYVTEHSHPEMYIFLC 66  
 DB 5 NQIVISQFLILGP--IPPEHWHLYTLLNLTIIILNLTIIILDSLHIMYIFLS 64  
 QY 67 LMSGIDILISTSSMPKMLAFWNSITIQDACLQIFAIHSLSGNESTVLLAMAFDRYV 126  
 DB 65 LKSLSPDSCFSVTPKILQNMNQDPSIPVADCLTQWFFLFGDLSFLVAMAYDRYV 124  
 QY 127 YVAICPLHATVLTLPRTKIGVA--AVRGAALMAPLVPFIKQLPFGCSNLSHSYCLH 183  
 DB 125 YVAICPLHATVLTLSMSPKFCVCSLSWPNVLYSMLHTLLARLSFCNDVIPEFCDIS 181  
 QY 184 QDVKLACDDIRVN----VYGLIVISAIGLSLISFSYLLIKTVLG--TREAOAK 237  
 DB 182 SALKLAFSTRNEMWVIFIMGLIIVT--PFLIIISVARIYSILTKVSSKIGICK 236  
 QY 238 AFGTGVSHVCAVFIFFYPIFGLSMWRHRSKRDSPLPILANITYLLVPVLANPIYGVK 296  
 DB 237 AFGTGVSHVCAVLEFYGTVIGLYLC--SSANSTLKTQVWAMNITVTPMLNPFYISL 293  
 QY 297 TKRIRQIRL 309  
 DB 294 NRDKKAISVH 306

RESULT 11  
 A37286  
 olfactory receptor 115 - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 26-Aug-1999  
 C/Accession: A37286  
 R/Buck, L.; Axel, R.  
 Cell 65, 175-187, 1991  
 A/Title: A novel multigene family may encode odorant receptors: a molecular basis for od  
 A/Reference number: A23701; MUID:91191556; PMID:1840504  
 A/Accession: A37286

A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-314 <BUC>  
 C:Cross-references: GB:M64392; NID:9205845; PID:AAA41755.1; PID:9205846  
 C:Superfamily: olfactory receptor OR14  
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 26.3%; Score 425.5; DB 2; Length 314;  
 Best Local Similarity 33.2%; Pred. No. 3,8e-30;

Matches 104; Conservative 70; Mismatches 114; Indels 25; Gaps 9;

QY 7 NESATYFLIGLPGLEBAQ--FWLAPLCSLYLAVIGNLTIYVTEHSLHEPMYIF 64  
 DB 5 NOTVISQFLFLPDPSEHGVYALF--LSKYLTVTGNLIIILHLSHHTPMYLF 62  
 QY 65 LCMLSGIDILISTSSMPKMLAFWENSTTIOPDACLQIFAHISLGSVESTVLAMAFDR 124  
 DB 63 LSNLSPSDLCFSSVMPKLLQNMOSQVPSIPAGCLTQLYFYLPADLESFLVAMAYDR 122  
 QY 125 YVAICHPLRHATVLTLPRTKIGVAAYVGAALM--APL-PVFIKQLPFGCSNLSHSY 180  
 DB 123 YVAICHPLRHATVLTLPRTKIGVAAYVGAALM--APL-PVFIKQLPFGCSNLSHSY 178  
 QY 181 CLHODVMKACDDIRNVVYGLVITSAIGDLSLISFSYLLIKTVGL-TREA 234  
 DB 179 CDSPLKLSGSDTHNNELVFMGSLVIVIPFV-----LIIVSARVVASILKVPISRG 233  
 QY 235 QAKARFTCVSHVCAVFIYVFPFGISWVRHSKRDSPLVILANTYLLVPPVNPVYVG 294  
 DB 234 HKISTCGSHLSVSLFYGIITGLYLP--SANNSTVKEVMMMYTVVPMNPFYIS 291  
 QY 295 VTKKEIRQRIRL 307  
 DB 292 LNRDMKEALIRV 304

## RESULT 12

S29707

Olfactory receptor OR5 - rat

C:Species: Rattus norvegicus (Norway rat)

C&gt;Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 26-Aug-1999

C:Accession: S29707; B37286

R:Ramung, K.; Krieger, J.; Strotmann, J.; Boekhoff, I.; Kubick, S.; Baumstark, C.; Breez

Nature 361, 353-356, 1999

A&gt;Title: Cloning and expression of odorant receptors.

A:Reference number: S29707; PMID:93149273; PMID:7678922

A:Accession: S29707

A:Molecule type: mRNA

A:Residues: 1-314 &lt;RAM&gt;

R:Buck, L.; Axel, R.

Cell 65, 175-187, 1991

A&gt;Title: A novel multigene family may encode odorant receptors: a molecular basis for od

A:Reference number: A23701; PMID:91191556; PMID:1840504

A:Accession: B37286

A&gt;Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 193-236 &lt;BUC&gt;

A:Cross-references: GB:M64375; NID:9205811; PID:AAA41738.1; PID:9205812

C:Superfamily: olfactory receptor OR14

C:Keywords: G protein-coupled receptor; membrane protein

Query Match 26.2%; Score 423.5; DB 2; Length 314;  
 Best Local Similarity 34.4%; Pred. No. 5.7e-30;

Matches 104; Conservative 64; Mismatches 119; Indels 15; Gaps 7;

QY 7 NESATYFLIGLPGLEBAQ--FWLAPLCSLYLAVIGNLTIYVTEHSLHEPMYIF 64  
 DB 5 NOTVISQFLFLPDPSEHGVYALF--LSKYLTVTGNLIIILHLSHHTPMYLF 62  
 QY 65 LCMLSGIDILISTSSMPKMLAFWENSTTIOPDACLQIFAHISLGSVESTVLAMAFDR 124  
 DB 63 LSNLSPSDLCFSSVMPKLLQNMOSQVPSIPAGCLTQLYFYLPADLESFLVAMAYDR 122

QY 125 YVAICHPLRHATVLTLPRTKIGVAAYVGAALM--APL-PVFIKQLPFGCSNLSHSY 180  
 DB 123 YVAICHPLRHATVLTLPRTKIGVAAYVGAALM--APL-PVFIKQLPFGCSNLSHSY 178  
 QY 181 CLHODVMKACDDIRNVVYGLVITSAIGDLSLISFSYLLIKTVGL-TREA 239  
 DB 179 CDSPLKLSGSDTHNNELVFMGSLVIVIPFV-----LIIVSARVVASILKVPISRG 238  
 QY 240 GTCVSHVCAVFIYVFPFGISWVRHSKRDSPLVILANTYLLVPPVNPVYVGKTE 299  
 DB 239 SSCGSHLSVSLFYGIITGLYLP--SANNSTVKEVMMMYTVVPMNPFYIS 296  
 QY 300 IR 301  
 DB 297 IR 298

## RESULT 13

H45774

Odorant receptor 3 - channel catfish

C:Species: Ictalurus punctatus (channel catfish)

C&gt;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 04-Sep-1998

C:Accession: H45774

R:Ngai, J.; Dowling, M.M.; Buck, L.; Axel, R.; Chess, A.

Cell 72, 657-666, 1993

A&gt;Title: The family of genes encoding odorant receptors in the channel catfish.

A:Reference number: H45774; PMID:93201590; PMID:7516654

A:Accession: H45774

A&gt;Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-321 &lt;NGA&gt;

A:Experimental source: olfactory epithelium

A&gt;Note: sequence extracted from NCBI backbone (NCBI:127745)

C:Superfamily: olfactory receptor OR14

C:Keywords: olfaction; transmembrane protein

Query Match 26.0%; Score 420.5; DB 2; Length 321;  
 Best Local Similarity 32.4%; Pred. No. 1.1e-29;

Matches 101; Conservative 69; Mismatches 127; Indels 15; Gaps 6;

QY 9 SSMATY-----FLIGLPGLEBAQFWLAPLCSLYLAVIGNLTIYVTEHSLHEPMYI 63  
 DB 14 TMTFIRPSTFYTGYNIPKAYYLF-LCPVYVTVTFGNSFTMGITVLARSHTAKYI 72  
 QY 64 FLCLMSGIDILISTSSMPKMLAFWENSTTIOPDACLQIFAHISLGSVESTVLAMAFDR 123  
 DB 73 AYFNALASDLCCSSALIPKLLDMLFENSGISYERACLSMFFVYCMTLQCTLLALAYD 132  
 QY 124 RYVAICHPLRHATVLTLPRTKIGVAAYVGAALM--APL-PVFIKQLPFGCSNLSHSY 183  
 DB 133 RLVAICHPLRHATVLTLPRTKIGVAAYVGAALM--APL-PVFIKQLPFGCSNLSHSY 192  
 QY 184 QDVMKACDDIRNVVYGLVITSAIGDLSLISFSYLLI-----LKTVGLTREDAQAKAF 239  
 DB 193 GRYLKLACDNDNTNSIMNGVCTATLLYFLILILISYVCTIGALAKIAGV---EYVKAM 249  
 QY 240 GTCVSHVCAVFIYVFPFGISWVRHSKRDSPLVILANTYLLVPPVNPVYVGKTE 299  
 DB 250 KCTTSHLILVAVFYLLIIVTVV-ALTTRIDNIRIINTALQTIIPMNPFIYVLTKEE 308  
 QY 300 IROIRILRF-HV 310  
 DB 309 VMQATLGLYKHI 320

## RESULT 14

S29711

Olfactory factor OR37 - rat

C:Species: Rattus norvegicus (Norway rat)

C&gt;Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 04-Sep-1998

C:Accession: S29711

R:Ramung, K.; Krieger, J.; Strotmann, J.; Boekhoff, I.; Kubick, S.; Baumstark, C.; Breez

A/Title: Cloning and expression of odorant receptors.  
 A/Reference number: S29707; MUID:93149273; PMID:7678922  
 A/Accession: S29711  
 A/Molecule type: mRNA  
 A/Residues: 1-305 <RAM>  
 C/Superfamily: olfactory receptor OR14

Query Match 25.9%; Score 418.5; DB 2; Length 305;  
 Best Local Similarity 32.5%; Pred. No. 1.5e-29;  
 Matches 100; Conservative 63; Mismatches 115; Indels 31; Gaps 5;

QY 15 ILILPELEAQLAPLPCSLYIAVLGNLTIIVYRTHSHHEPMYIFLCMLSGIDIL 74  
 DB 1 LILGLSGYPTKEILYFVIVLMYLVHTGNGVLIISIFDSHLHTEMFEFLGLSLFDIC 60  
 QY 75 ISTSMPEKMLAIFWNSGTIQFDCLIQIFAIHSLGMESTVLLAAPPVYALICPLRH 134  
 DB 61 YTTSSVSTLVSLSKRNTSFGCTVQMFVGFAMSTECLELGMVAFDRVAILCPLEKY 120  
 QY 135 ATVILTPVTKI-----GVAAYVRGAALMAPLPVFIKQLPFCRSNITLSHYCLHGV 186  
 DB 121 SVIMSKREVYVMSASWFSGINSVQTSIAM-----RLPFCGNVYNHPTCEVLAV 172  
 QY 187 MKLACDDIRVNVVYGLVITISAGLSLISPSYLLIKTVLGL-TREAQAKAFGTCVSH 245  
 DB 173 LKLAGADISLNTVMTVSNAPFLVPLLFISYVILTYTILMNSAGRRKAFSTCSAH 232  
 QY 246 VCAVFIYVPEIGLSMVHFRSKRDSPL-----PVLANIYLLVPPVLPVPIYGVK 296  
 DB 233 LTVVVIYFGTFSM-----YAKPKSODLTGDKFQTSDKISLFYGVVTPMLNPIIYSR 287  
 QY 297 TKEIRORI 304  
 DB 288 NKDVKAAY 295

RESULT 15  
 S51356  
 olfactory receptor - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C/Accession: S51356; S47014  
 R/Gat, U.; Nekrasova, E.; Lancet, D.; Natochin, M.  
 Eur. J. Biochem. 225, 1157-1168, 1994  
 A/Title: Olfactory receptor proteins. Expression, characterization and partial purification  
 A/Reference number: S51356; MUID:95045546; PMID:7957207  
 A/Accession: S51356  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-309 <GAT>  
 A/Cross-references: EMBL:X80671; NID:9517365; PIDD:CAAS6697.1; PID:9517366  
 R/Gat, U.; Nekrasova, E.; Lancet, D.; Natochin, M.  
 submitted to the EMBL Data Library, July 1994  
 A/Description: Olfactory receptor proteins: expression, characterization and partial purification  
 A/Reference number: S47014  
 A/Accession: S47014  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-309 <GA2>  
 A/Cross-references: EMBL:X80671; NID:9517365; PIDD:CAAS6697.1; PID:9517366  
 C/Superfamily: olfactory receptor OR14

Query Match 25.9%; Score 418.5; DB 1; Length 309;  
 Best Local Similarity 33.0%; Pred. No. 1.6e-29;  
 Matches 104; Conservative 63; Mismatches 119; Indels 29; Gaps 8;

QY 6 GNSSATYFLLIGLPEAQLAPLPCSLYIAVLGNLTIIVYRTHSHHEPMYIFL 65  
 DB 5 GNSAVVVFVLTGTPKPELLPLFPLFLVIVLVTVGNLMLIIVSPILHTPMYIFL 64  
 QY 66 CMLSGIDILISMPMTAIFWNSGTIQFDCLIQIFAIHSLGMESTVLLAAPPVY 125  
 DB 65 SSLSFVDLCYSTVTEKMLVNFGLKGNFTYSECMAQFFFPALFVTEGYLLTVMAVDYR 124

QY 126 VAICPELHAIVTLTPVTKIY-----AAVVRGAALMAPLPVFIKQLPFCRSNITL 176  
 DB 125 VAICPELHYVIMG-SRISLVLVAFSLGLFSAVHTSMM-----NLSFCKSYII 175  
 QY 177 SHSYCLHGVKMLACDDIRVNVVYGLVITISAGL-LDSLISPSYLLIKTVLGL-TRE 233  
 DB 176 SHYFDDALPLKLACNSVHNLNEL--LFIIGLANTVPTLAVASTVFIKSIIRISSE 233  
 QY 234 AQAQAFGTCVSHVCAVPIFY--VPIGLSMVHFRSKRDSPLPVLANIYLLVPPVLP 291  
 DB 234 GRSYAFGTCSSHLMVAGIFFGSIITFMYLKSSNSLQCEK-----VSSVFYTTVTPMLNPL 289  
 QY 292 VYGVKTEIRORILR 306  
 DB 290 IYSLRNQVKAALGR 304

Search completed: February 9, 2004, 16:25:43  
 Job time : 21 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 16:05:17 ; Search time 163 Seconds

(Without alignments)  
8491.887 Million cell updates/sec

Title: US-10-017-066a-1

Perfect score: 3136  
Sequence: 1 cagagagcgtatcttcagt.....aaaaaaaaaaaaaaaa 3136

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: 1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	377.4	12.0	963	US-09-439-313-526	Sequence 526, App
2	372.6	11.9	1474	US-08-465-980-1	Sequence 1, Appl
3	372.6	11.9	1474	US-09-053-303-1	Sequence 1, Appl
4	372.6	11.9	1474	US-09-339-115-1	Sequence 1, Appl
5	372.6	11.9	1474	PCT-US95-07093-1	Sequence 1, Appl
6	263	8.4	1539	US-09-668-680-13	Sequence 13, Appl
7	215.4	6.9	231	US-09-439-313-462	Sequence 462, App
8	215.4	6.9	231	US-09-352-616A-462	Sequence 462, App
9	213.4	6.8	234	US-09-016-434-1047	Sequence 1047, App
10	134	4.3	1854	US-09-016-434-1312	Sequence 1312, App
11	125	4.0	1990	US-09-016-434-1056	Sequence 1056, App
12	121.4	3.9	1713	US-08-467-947A-1	Sequence 1, Appl
13	121.4	3.9	1713	US-08-467-947A-1	Sequence 1, Appl
14	120.2	3.8	3459	US-09-016-434-1363	Sequence 1363, App
15	112.6	3.6	251	US-09-439-313-352	Sequence 352, App
16	112.6	3.6	1062	US-09-353-616A-352	Sequence 352, App
17	108.2	3.5	1080	US-09-668-680-10	Sequence 10, Appl
18	104.4	3.3	1080	US-09-668-680-9	Sequence 9, Appl
19	104	3.3	951	US-09-465-901-17	Sequence 17, Appl
20	101.2	3.2	951	US-09-465-901-47	Sequence 47, Appl
21	100.6	3.2	1297	US-09-668-680-11	Sequence 11, Appl
22	95	3.1	945	US-09-016-434-1114	Sequence 1114, App
23	93.2	3.0	1282	US-09-016-434-1413	Sequence 1413, App
24	91.6	2.9	542	US-09-016-434-627	Sequence 627, App
25	86.6	2.8	966	US-08-748-506-8	Sequence 8, Appl
26	86.4	2.8	669	US-09-465-901-25	Sequence 25, Appl
27	83.2	2.7	1828	US-08-988-876-2	Sequence 2, Appl

28	82.8	2.6	1438	US-09-016-434-1313	Sequence 1313, App
29	82	2.6	966	US-09-016-434-1372	Sequence 1372, App
30	81.8	2.6	945	US-08-748-506-7	Sequence 7, Appl
31	81.6	2.6	952	US-09-016-434-1115	Sequence 1115, App
32	79.8	2.5	966	US-08-748-506-5	Sequence 5, Appl
33	75	2.4	669	US-09-465-901-11	Sequence 11, Appl
34	73.4	2.3	966	US-08-748-506-6	Sequence 6, Appl
35	73	2.3	669	US-09-465-901-13	Sequence 13, Appl
36	72.4	2.3	270	US-09-016-434-150	Sequence 150, App
37	71.2	2.3	666	US-09-465-901-13	Sequence 13, Appl
38	70	2.2	648	US-09-016-434-1375	Sequence 1375, App
39	69.8	2.2	669	US-09-465-901-41	Sequence 41, Appl
40	69	2.2	681	US-09-465-901-29	Sequence 29, Appl
41	67.8	2.2	1525	US-09-157-603-4	Sequence 4, Appl
42	67.8	2.2	1525	US-09-587-436-4	Sequence 4, Appl
43	67.8	2.2	1525	US-08-927-165A-4	Sequence 4, Appl
44	67.4	2.1	3238	US-08-123-934A-5	Sequence 5, Appl
45	67.4	2.1	3238	PCT-US94-10080-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-09-439-313-526  
Sequence 526, Application US/09439313  
Patent No. 6329505  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, David C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang, Yugu  
APPLICANT: Reed, Steven G.  
APPLICANT: Kalos, Michael  
APPLICANT: Fanger, Gary  
APPLICANT: Retter, Mark  
APPLICANT: Solk, John  
APPLICANT: Day, Craig  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C9  
CURRENT APPLICATION NUMBER: US/09/439,313  
CURRENT FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 526  
LENGTH: 963  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-439-313-526  
Query Match 12.0%; Score 377.4; DB 4; Length 963;  
Best Local Similarity 64.8%; Pred. No. 4,3e-82;  
Matches 577; Conservative 0; Mismatches 311; Indels 3; Gaps 1;  
171 CTTCATCTCAATAGAGCCCTGCTTTAGAGAGGCTAGTTCTGATGGCCCTCCATT 230  
30 CTTTGCGCTTAGTATCCAGATAGAGAAACCATTTCTGGTGGTCCCT 89  
231 GTGCTCCCTTACCTTATTTGCTGCTAGAGTACATTCATCTACATTTGCGGAC 290  
90 CTTTCCAGTATGATGAGGCAATTTTGAAGATCGATCGTCTTCATCGTAGAGC 149  
291 TAGACAGCCCTGAGAGCCCATATATATTTTTCAGATGCTTTCAGGATGACAT 350  
150 GGAAGCAGCCCTGAGAGCCCATATATCTTTCTGAGATGCTTTCAGGATGACAT 209  
351 CCTCATCTCACTTATCATTCATCCCAATATGCTGCGCATCTTCTGCTTCACTAC 410  
210 GGCCTTATCCATCCATCCATGCTTATCTTCCCTTTCTGATTTGATTTCCGAGA 269  
411 CATCAGTTTATGCTTGTCTACAGATTTTTCATCCATCTTATCTGAGATGAGA 470

Db	270	GATTAGCTTTGAGGCGCTGTCTTACCCAGATGTTCTTTATTCATGCGCCCTTCACGCAATTGA	329
Qy	471	ATCCACAGAGTGCCTGCTGGCCATGAGCTTTTGAACCGCATATGGCCATCTGTACCCACATGGG	530
Db	330	ATCCACCATCTGTGTGGCCATGAGCCCTTTGACCGCTTATGTGGCCATCTGCAACCACTGGG	389
Qy	531	CCATGCCACAGTACTTACCTTACCTGTGCTGTGTACCCAAATTGGTGTGCTGTGTGTGGG	590
Db	390	CCATGTGCGAGTGCCTCAACATACAGTAAACAGCCAGATTGGATCTGTGTGTGTGG	449
Qy	591	GGGGGCTGCACATGATGAGCAACCCCTTCCTGTCTTCAATCAAGAGATGGCCCTTGGCCGCTC	650
Db	450	CGGATCCCTCTTTTTTTTCCTCCATGCTCTGTGTATCAAGGGCTGGGCTTCTGCACTTC	509
Qy	651	CAATATCTTTTCCCATCTTACTGCTTACACCAAGATGTATGAAGCTGGCTGTATGA	710
Db	510	CAATGTCTCTGTGCACCTCTTATGTGTGCACCAAGATGTATGAAGTGGCTATGACAG	569
Qy	711	TATCCGGGTCAATGTGTGTATGTGGCTTATGTGTATATCTCCGCAATGAGCTGTGACATC	770
Db	570	CATTTTGGCCCAATGTGTATGTGTCTTACGTCAATTTCTGTGTATGAGGCTGTGACGT	629
Qy	771	ACTTCTCATCTCTCTCTCAATCTGTCTTATTTCTTAAAGCTGTGTGGCTTGAACAGTGA	830
Db	630	AATGTTCAATCTCTGTCTCTTATTTCTATATACGAACGGTTCTGCATGCTCTTCA	689
Qy	831	---AGCCACAGGCGCAAGGCATTTGGGCACTGTGGCTGTCAATGTGTGTCTGTTCATAT	887
Db	690	GTCAGAGCGGGCCAAAGCCTTTGGAACCTGTGTCAACATTTGTGTATCTGCTCTT	749
Qy	888	CTATGTACTCTTTCATTTGATTTTCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	947
Db	750	CTATGTGCACTTATGT	809
Qy	948	GCTGCCGCTCATCTTGGCCCAATATCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1007
Db	810	TGTGCTGT	869
Qy	1008	CTATGTGATGTGAACAAAGAGATTCGACAGCGCATCTTGCATTTTCCA	1058
Db	870	CTATGTGTGCAAAACCAACCAATGACGAACAGGAGCTGTGTGTGTGTGTGTGTGT	920

RESULT 2  
 US-08-465-980-1  
 Sequence 1, Application US/08465980  
 Patent No. 5756309  
 GENERAL INFORMATION:  
 APPLICANT: Soppet, Daniel R.  
 APPLICANT: Li, Yi  
 APPLICANT: Rosen, Craig A.  
 APPLICANT: Ruben, Steven M.  
 TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAL70  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
 ADDRESSEE: STUART & OLSTEIN  
 STREET: 6 Becker Farm Road  
 City: Roseland  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07068  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/465,980  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:

	NAME:	Ferraro, Gregory D.
	REGISTRATION NUMBER: <td>36,134</td>	36,134
	REFERENCE/DOCKET NUMBER: <td>325800-446</td>	325800-446
	TELECOMMUNICATION INFORMATION: <td></td>	
	TELEPHONE: <td>201-994-1700</td>	201-994-1700
	TELEFAX: <td>201-994-1744</td>	201-994-1744
	INFORMATION FOR SEQ ID NO: <td>1</td>	1
	SEQUENCE CHARACTERISTICS: <td></td>	
	LENGTH: <td>1474 base pairs</td>	1474 base pairs
	TYPE: <td>nucleic acid</td>	nucleic acid
	STRANDEDNESS: <td>single</td>	single
	TOPOLOGY: <td>linear</td>	linear
	MOLECULE TYPE: <td>CDNA</td>	CDNA
	FEATURE: <td></td>	
	NAME/KEY: <td>CDS</td>	CDS
	LOCATION: <td>274..1233</td>	274..1233
	US-08-465-980-1	
Query Match	11.9%	Score 372.6; DB 1; Length 1474;
Best Local Similarity	64.4%	Pred. No. 7.7e-81;
Matches 574; Conservative	0; Mismatches	314; Indels 3; Gaps 1
QY	171	CTTCATCTCTAATAGCCCTCCCTGGTTTAAAGAGAGCTCAGTTCTGGTGGCCCTTCCAT
DB	303	CTGTGTGCTATTGTGTATCCAGGATTAGAAAGCCCATTTCTGGTGGCTTCCCT
QY	231	GTGCTCCCTCCTACCTTATTTGTGTGCTAGGTAACCTGACATTCATCTCATTTGGGAG
DB	363	CCTTTCCATGTATGTAGTGGCAATGTGGAACCTGCATCGTGGCTTTCATGTGTAAGAC
QY	291	TGAGCAGAGCCCTGCATGAGCCCATGTATATATTTCTTTGATGCTTTGAGGATTTGACAT
DB	423	GGAAGCAGAGCTGCACAGCGTCGATGTACCTCTTCTGTGCATGCTTGACGCATTTGACCT
QY	351	CCTGATCCCAACCTCATTCCATGGCCCAAAAGCTGGGCACTCTTGTGTTCAATCCACATAC
DB	483	GGCCTTATCCATCATCCACATGACCTTAAGATCTTGCCCTTTTGTGGTTGATTTCCCGAGA
QY	411	CATCCAGTTTGATGCTGTGTGTGTACAGATTTTGGCATCCATCCATCTTATCTGATGATGA
DB	543	GATTAGCATTTAGAGCGCTGTCTTTACCCAGATGTCTTTATTCATGCGCCCTCAGCATTTGA
QY	471	ATCCAGAGTGTGTGTGGCCATAGGCTTTTGAACCGGTATGTGGCCATTCGTCAACCATGGG
DB	603	ATCCACATCTCTGTGTGGCCATAGGCTTTTGAACCGGTATGTGGCCATTCGTCAACCATGGG
QY	531	CCATGCCAAGTACTTACGTTGCGCTGTGTGTACCCAAATTGGTGTGCTGTGTGGTGGC
DB	663	CCATGTGCGAGTGTGTCAAAATACAGTAACAGCCCAAGTTGGCATCTGTGGTGTGGTGGC
QY	591	GGGGGCGTGCATGATGAGCACCCCTTCCGTCGTCCTTATCAAGAGAGGCTGCCCTCTGGCGGTC
DB	723	CGGATCCCTCTTTTCTTCCCATGCTCTGTGTGTATCAAGAGGCTGCCCTTCTGTGCATTC
QY	651	CAATATCTCTTCCATTTCTTACTGTGCTACACCAAGATGTCTAGTGAAGCTGTGCTGTATGA
DB	783	CAATGTCTCTCGACATCTTATGTGTGTCCACCAAGATGTATGAAGTTAGGCTGTATGACGA
QY	711	TATCCGGGTCAATGTGTGTATAGGCTTATGTGTATCATCTCCGCAATTTGGCTGTGATCTC
DB	843	CACCTTGGCCATGTGTGTATGTGTCTTATCTGTCAATCTGTGTGTATGAGGCGTGTGACGT
QY	771	ACTTCTCATCTCTTCTCATATCTGTCTTATTTCTTAAAGCTGTGTGTGTGTGACACGTGA
DB	903	AATTTTCATCTCTCTTGTCTTCTTATTTCTGATATATAGAACGGTGTCTGCAACCTGTCCAA
QY	831	---AGCCAGAGCCAGGACATTTGGCATTGGGCTTCCATCTGTGTGTGTGTGTGTGTGTGT
DB	963	GTCAGAGCGGGCCAGGCGCTTTGGAACCTGTGTGTACACATTTGTGTGTGTGTGTGTGTGT
QY	888	CTATGTACTTTTCAATTGGATTGTTCACATGTGTGCATGCTGCTTTAGCAAGCGGCTGTACTCTC
DB	1023	CTATGTGCACCTTATTTGGCGCTCTCAGTTGTACACCGGCTTTGGAAACAGCCCTTACATTC



```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/339,115
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/053,303
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-446
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1474 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 274..1233
US-09-339-115-1

Query Match      11.9%; Score 372.6; DB 4; Length 1474;
Best Local Similarity 64.4%; Pred. No. 7.7e-81;
Matches 574; Conservative 0; Mismatches 314; Indels 3; Gaps 1;

QY 171 CTTCATCCCTACCTTATGCTGTGCTAGGTAATTCATCATCTTATGCTGCTTCCAT 230
DB 303 CTGTGTGCTTATGCTATCCAGATTAGAAAGCCATTCTGTGGTTGCTTCCCT 362
QY 231 GTGCTCCCTACCTTATGCTGTGCTAGGTAATTCATCATCTTATGCTGCTGAGC 290
DB 363 CTTTCCATGTATGCTAGTGTGCTAGTGTGAAAGTGTGCTTCTTATGCTGTAAGAC 422
QY 291 TGAGCAGAGCTTGCATGAGCCCATGTATATTTCTTGCATGCTTTCAGGCAATGACAT 350
DB 423 GGAAGGAGAGCTTGCAGCTCCGATGTACCTCTTCTGCTGATGCTTGCAGCCATTGACCT 482
QY 351 CCTCATCTCAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
DB 483 GGCCTTATCCATCCATCCATGCTTAAAGATCCTGCTGCTTCTGCTGCTTCTGCTGCTGCT 542
QY 411 CATTCAGTTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 470
DB 543 GATTAGCATTTGAGGCTGCTGCTTACCCAGATGCTTATTCATGCTTCCCTTCAGCCATGGA 602
QY 471 ATCCAGAGTCTGCTGCTGCTGCTTTCAGGCTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530
DB 603 ATCCAGATCTGCTGCTGCTGCTTTCAGGCTTTCAGGCTTTCAGGCTTTCAGGCTTTCAGGCT 662
QY 531 CCAATGCCAGTACTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
DB 663 CCATGCTGAGTGTCTCAACAATACAGTAACAGCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
QY 591 GGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
DB 723 CGGATCCCTTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
QY 651 CAAATATCCCTTCCATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 710
DB 783 CAAATGCTCTCCGACCTGCTATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842
QY 711 TATCCGGGTAAATGTGTGTATGCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
DB 843 CACTTTCGCCAATGTGTATGTGTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902

```

```

QY 771 ACTTCATCTCTCTCTCATATCTGCTTATCTTAAAGCTGTGTGGCTTGAACAGTGA 830
DB 903 AAGTTCATCTCTCTCTCTCATATCTTCTGATTAATACAGAGGTTCTGCAACTGCTTCCAA 962
QY 831 ---AGCCAGAGCCCAAGGCAATTTGGCACTTGGCTCTCTCATGTGTGTGCTGTCTCATATT 887
DB 963 GTGAGAGCGGCGCAAGGCTTTTGAACCTGTGTGTACACATTTGTGTGTACTGCTT 1022
QY 888 CTATGTACCTTTCATTTGATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 947
DB 1023 CTATGTGCTCATTTATGCTGCTCTCAGTTGTGACACCGCTTTGAAAACAGCCTTATCCCAT 1082
QY 948 GCTGCGCGCTATCTTGGCCCATATCTATCTGTGTGCTTCTCTGCTGCTGCTGCTGCTGCTGCT 1007
DB 1083 TGTGCGGTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
QY 1008 CTATGAGTGAAGACAAGAGATTTGACAGCGCATCTTGTGACTTTTCCA 1058
DB 1143 CTATGTGCTCAAAACCAACAGATCAGAACACGGGTGTGCTGCTATGTTCA 1193

```

## RESULT 5

```

PCT-US95-07093-1
Sequence 1, Application PC/TUS9507093
GENERAL INFORMATION:

```

```

APPLICANT: Soppet, Daniel R.
APPLICANT: Li, Yi
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAL70
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
ADDRESSER: STUART & OLSTEIN
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07093
FILING DATE:
CLASSIFICATION:

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,980
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-446
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1474 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 274..1233
PCT-US95-07093-1

```

```

Query Match      11.9%; Score 372.6; DB 5; Length 1474;
Best Local Similarity 64.4%; Pred. No. 7.7e-81;
Matches 574; Conservative 0; Mismatches 314; Indels 3; Gaps 1;

```

```

QY 171 CTTCATCCCTAATAGAGCCCTCCCTGGTTTGAAGAGGCTCAGTTCGTGGCCCTCCATT 230
Db 303 CTGTGTGCTTATGTGTATCCAGAGATAGAGAAAGCCATTTCGTGGTGGCTTCCCTCCT 362
QY 221 GTGCTCCCTTACCTTATTTGCTGTGCTAGTAAGTACATCATCTACATTTGGCGGAC 290
Db 363 CTTTCCAGATATGATGAGGGAATGTGGAAGATGCTGTGTCTTCACTGTAAGAGAC 422
QY 291 TGAGCAGAGCCCTGATGAGCCCATGTATATTTCTTTCATGCTTCAGGATTGACAT 350
Db 423 GGAAGAGAGCCCTGACAGCGCTCCGATTAACCTCTTCTCTGATGCTTGAAGCATTAAGCT 482
QY 351 CCTCATCTCCACCTTATCATTCATGCCAAATGCTGCCATCTTCTGTTCATTTCCACTAC 410
Db 483 GGCCTTATCCACATCCACCATGCTCCATGATCTTCCCTTTCTGTGTGATTCCTCCAGA 542
QY 411 CATCCAGTTTATGCTGTGCTGCTACAGATTTTGGCCATCCATCCCTTATCTGGCATGA 470
Db 543 GATTAGCATTTAGAGCCCTGCTTACCCAGATGTTCTTTATATAGCCCTCTTACGCAATTGA 602
QY 471 ATCCACAGTCTGTGGCCATGAGCTTTTGAACCGCTATGTGCCATCTGTCAACCCACTGCG 530
Db 603 ATCCACATCTGCTGGCCATGAGCTTTGACCGTATATGTGCCATCTGCCACCCACTGCG 662
QY 531 CCATGCCACAGTACTTACGTTGCCCTGCTGCTACCCAAATGTGTGTGCTGCTGTGTGCG 590
Db 663 CCATGCTGAGTGTCAACATACATACAGTAAACGCCCATGTTGCAATGCTGTGTGCG 722
QY 591 GGGGGCTGACATGATGAGCAACCCCTTCTGCTTCACTACAGAGTGGCCCTTCTGCGCTC 650
Db 723 CGGATCCCTCTTTTTCCTCCAGCTGCTGCTGATCAAGGGCTGTGCTTCTGCGCTC 782
QY 651 CAATATCTTCTTCCATCTCTTCTGCTGCTTCAACCAAGATGTCAATGAGTGGCTGTGATGA 710
Db 783 CAATGTCTCTGCACTCTCTTATGTGTCCACAGATGTATGAGTGTGCTATGAGCA 842
QY 711 TATCCGGGTCAATGTCTGTATGAGCTTATGCTATCATCTCCGCAATGGCCCTGAGCTC 770
Db 843 CACTTGGCCCATGTGTGTATGTGTATGTCTTACGCACTTGTGTGTGTGTGTGTGTGTGT 902
QY 771 ACTTGTCACTCTCTTCTATCTATCTGTCTTATCTTAAAGCTGTGTGTGTGTGTGTGTGT 830
Db 903 AATGTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 962
QY 831 ---AGCCAGAGCCAGAGATTTGGCACTTGGCTGTCTGTATGTGTGTGTGTGTGTGTGT 887
Db 963 GTCAAGAGCGGCGCAAGGCTTGTGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
QY 888 CTATGTACCTTTCATTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 947
Db 1023 CTATGTGCACTTATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1082
QY 948 GCTGCGCCGCTATCTTGGCCATATCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1007
Db 1083 TGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1142
QY 1008 CTATGAGTGAAGACAAAGAGATTCAGACAGGCACTCTTCACTTTTCCA 1058
Db 1143 CTATGCTGCAAAACCAACAGATCAGAACAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1193

```

RESULT 6  
US-09-668-680-13  
Sequence 13, Application US/09668680

Patent No. 6436703  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Zhou, Ping  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Wang, Jian-Rui

```

APPLICANT: Xue, Aidong J.
APPLICANT: Xu, Chongjun
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. 6436703el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 790CIP2A
CURRENT APPLICATION NUMBER: US/09/668,680
CURRENT FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ IDS: 13
SOFTWARE: pl_genes Version 2.0
SEQ ID NO: 13
LENGTH: 1539
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (130)..(1539)
US-09-668-680-13

```

Query Match 8.4%; Score 263; DB 4; Length 1539;  
Best Local Similarity 57.6%; Pred. No. 38-54;  
Matches 511; Conservative 0; Mismatches 370; Indels 6; Gaps 2;

```

QY 183 AGGCTCCCTGTTTGAAGAGGCTCAGTTCGTGTGGCTTCCATTTGCTCCCTCTTA 242
Db 636 AGGATCCAGAGGCTGAGAGGCGCCACTTCTGATGTGCATCCCTTCTGTGCCATGTA 695
QY 243 CCTTATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
Db 696 TCTTGTAGACCTGTGTGAAATGCTGCTCTATCTGCTGATGCTGATGCTGATGCTGATGCT 755
QY 303 GATGAGCCCATGTATATATTTCTTTGATGCTTTCAAGCATGACATCTCATCTCAC 362
Db 756 TATGACACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 815
QY 363 CTCATCCATGCCCAAAATGCTGCCCATCTTCTGTGTCAATTCATCATCCATGCAATTTGA 422
Db 816 TACCATGTGCCCAAGATGCTGCCCATTTGTGCTCCATGCTGATGATGATGATGATGATGAT 875
QY 423 TGTCTGTCTGATACATTTTTCATCCATCTCTTATCTGTGATGATGATGATGATGATGATGAT 482
Db 876 TGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 935
QY 483 GCTGCGCATGCTTTTGAACCGCTATGTGSCATCTGTACCCATGCGCATGCGCATGCT 542
Db 936 ACTTGTGCAATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 995
QY 543 ACTTACGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
Db 996 TCTCAACATGCTGTGTATAGGCAATTTGCTTTGTGTGCTATTCCTGATGATGATGATGAT 1055
QY 603 GATGAGACCCCTTCTGCTTATCATCAAGCACTGCTGCTTGTGCGCTCAATATCTTTTC 662
Db 1056 TGTCTCTCCCTTCTATCTTCTTCTGAGGCACTCCCTTCTATGCTGCTGCTGCTGCTGCT 1115
QY 663 CCATTCCTATGCTGCTACACCAAGATGCTATGAGTGGCTGCTGATGATGATGATGATGATGAT 722
Db 1116 ACACATATCTGTAGACATATGAGCATGCGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1175
QY 723 TGTGCTATAGCTTATGCTATCATCTGCGCCATTTGCTGAGCTGCTGCTGCTGCTGCTGCT 782
Db 1176 TATGTCTATGAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1235
QY 783 CTTCATATGCTGCTATTTTAAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
Db 1236 CATTTCTATGCTTTATCTTCAATGCAATCTTTTCACTTCACTGATGATGATGATGATGAT 1295
QY 840 CAGGCAATTTGCACTTGTGCTGTCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899

```



INFORMATION FOR SEQ ID NO: 1047:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 234 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: KIDNUT01  
CLONE: 998550  
US-09-016-434-1047

Query Match 6.8%; Score 213.4; DB 4; Length 234;  
Best Local Similarity 96.4%; Pred. No. 1.3e-42;  
Matches 214; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 334 CTTTCAGGATTCAGATCTCTATCTCCACCTCATCCATGCCCCAAATGCTGGCATCTC 393  
DB 1 CTTTCAGGATTCAGATCTCTATCTCCACCTCATCCATGCCCCAAATGCTGGCATCTC 60  
QY 394 TGGTTCATTCCTCACTACCATCCAGTTTATGCTTGTCTGCTACAGATTTTGGCATCCAC 453  
DB 61 TGGTTCATTCCTCACTACCATCCAGTTTATGCTTGTCTGCTACAGATTTTGGCATCCAC 120  
QY 454 TCCCTATCTGCGATGGAATCCACAGTCTGCTGGCCATTTTGACCGCTATGTGACC 513  
DB 121 TCCCTATCTGCGATGGAATCCACAGTCTGCTGGCCATTTTGACCGCTATGTGACC 180  
QY 514 ATCTGTACCCACCTGCGCCATGCGACAGTACTTACCTGCT 555  
DB 181 ATCTGTACCCACCTGCGCCATGCGACAGTACTTACCTGCT 222

## RESULT 10

US-09-016-434-1312  
Sequence 1312, Application US/09016434  
Patent No. 6500938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016.434  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 845-4166  
TELEFAX: (650) 845-0555  
INFORMATION FOR SEQ ID NO: 1312:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1854 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: 932085  
US-09-016-434-1312

Query Match 4.3%; Score 134; DB 4; Length 1854;  
Best Local Similarity 48.6%; Pred. No. 6.3e-23;  
Matches 430; Conservative 0; Mismatches 445; Indels 9; Gaps 2;

QY 151 AATGATTCAGGCTCATCTTCTCAATTAAGGCTCCCTGGTTAGAAAGGTGAC 210  
DB 332 AATCAACAGCATCTCTCACTCTCCCTGCGGCTGGCCATCCACAGAGAGCA 391  
QY 211 TTCTGTTGGCTTCCCATTTGCTCCCTCACTTATGCTGTGAGTACTTGACA 270  
DB 392 AACCTGTCTATGCTCCCTGTTCTTGGCATGTAATTTACACCCCTCTGGAGACCTCCTC 451  
QY 271 ATCATCTACATTTGCGGAGCTAGAGCAAGCCGTGATGAGCCCATGTATATTTCTTGC 330  
DB 452 ATCATTTGCTCAATTCAGTGAATCCCATCCACACGCTTATGTAATTTTCTCAGC 511  
QY 331 ATGCTTTCAGGATTCAGATTCCTCATCTCCACCTCATCCATGCCCCAAATGCTGGCATC 390  
DB 512 AACTTGTCTCTCTGACCTCTGCTTCTCTTCCGAGACCTATCCAGTTGTACAGAAC 571  
QY 391 TTCTGTTCAATTCATCCATCCATCCAGTTGATGCTTGTCTGCTACAGATTTTGGCATC 450  
DB 572 ATGCAAGAACAGAGACCATCCATCCCTTATGGGACTGCTGACCAATGTAATCTTTC 631  
QY 451 CACTCCTTATTCGAGATGGAATCCACAGTGCCTGGCCATGCTTTGACCGCTATGTC 510  
DB 632 CTGTTATTTGAGACCTGAGAGCTTCTCTTGTGGCCATGACCTATAGCGTAATGTC 691  
QY 511 GCCATCTGACCCACTGCGCCATGCGACAGTACTTACCTTGCCTGTGTCACCAAAAT 570  
DB 692 GCCATCTGCTCCCTGACACTACACCGCATCATAGAGCCCATGCTGTCTGCGCCCTG 751  
QY 571 GGTGTGCTGCTGTGTGGGGGGGGGTGACATGATGAGCAACCCCTTCCTTCAACAG 630  
DB 752 GTGGCGCTGTCCGGTGTGACCACTTCACATGCCATTTACACTTACATGAGCC 811  
QY 631 CAGTCCCTTCTGCGCGCTCCAAATCCCTTCCATTTCTTACTGCTACACCAAGATGTC 690  
DB 812 AGGTGTGTTTGTGACAGACATGATGCCCATCTTTCTGTGATATGTCTGCTG 871  
QY 691 ATGAAGCTGCGCTGTGATATATCCGGTCAATGTCGTATGAGCCTTATGCTCATATC 750  
DB 872 CTGAAGCTGCGCTTCTGTGACACTGAGTATGATGAGTGATATTTATCATGAGAGG 931  
QY 751 TCCGCAATTTGGCTGAGTCACTTCTCATCTCTCTCATCTGCTTATCTTAAGACT 810  
DB 932 CTGATCTTGTGATCCCATCTTCTTACTCATCTTGGGTCCTATGCAAGATGTCTCCTC 991  
QY 811 GTGTT--GAGCTTGAACAGTAAAGCCAGGCGCAAGGCTTTGGCATCTTGTGCTCTAT 867  
DB 992 ATCTCTAAGAGTCCCTTCTCTAAGGATATGCAAGGCTTCTCTAATTTGTGCTCCAC 1051  
QY 868 GGTGTGCTGTGTATATTTATGTAATCTTCAATGATGATGCAATGATGATGATGATG 927  
DB 1052 CTGTCTGTGTGTACTGTCTTATGTAACCGTATTTGCTCTACTT-----ATGCTTA 1105  
QY 928 AGCAAGGCGGTGATCTTCCGCTCCCGCTCATCTTGGCCATATCTATCTGTGCTTCT 987  
DB 1106 TCAGTAAATAGTTTACTTAAGAGCACTGTCAATGCTATGATGACATCTGTGTGACC 1165  
QY 988 CCTGTGCTCAACCAATTTGCTATGAGTGAAGAGCAAGAGAGAT 1031  
DB 1166 CCAATGCTAACCCTTATCTTACAGCTGAGGAACAGAGACAT 1209

## RESULT 11





CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/04079  
 FILING DATE: 30-MAR-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: STEFFE, ERIC K.  
 REGISTRATION NUMBER: 36,688  
 REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2540  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1713 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: both  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 116..1003  
 US-08-467-947A-1

Query Match 3.9%; Score 121.4; DB 2; Length 1713;  
 Best Local Similarity 48.0%; Pred. No. 6.9e-20;  
 Matches 378; Conservative 0; Mismatches 406; Indels 3; Gaps 1;

151 AATGAATCCAGTGTACATCTTCAATGAGGCTCCCTGTTTGAAGAGGCTCAG 210  
 137 AATCAGACAAATGTGACAGAGTCTCTACTGAGGATTTCTCCGAGCCCAAGATTGAG 196  
 211 TTCTGTTGGCCCTTCCATGTCGCTCCCTACCTTATGCTGCTGAGTACATGACA 270  
 197 ATGCTCTCTTTGGGCTCTTCTCCCTGTTTATGTTCTTCACTCCCTGGAAGAGGACC 256  
 271 ATCATCTACATTTGGGAGTACAGACAGCCCTGATGAGCCCATATATATTTCTTTC 330  
 257 ATCTGAGGGCTCATCTGACTGAGTCCAGACTCCAGCCCATGATCTTCTCTCTCA 316  
 331 ATGCTTTAGGATGAGTACCTCTCATCTCCACTCATCCAGCCCAAAATCTGGCCATC 390  
 317 CACCTGGCCGTTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376  
 391 TTCTGTTCAATTCACATCTACATCCAGTTTGAATGCTGCTGCTGCTGCTGCTGCTG 450  
 377 CTCTGATCCAGCCAGCCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436  
 451 CACTCTTATCTGGGATGAGTACCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 510  
 437 TTGAGTTTGGACATATGAAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 496  
 511 GGCATCTGTACCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 570  
 497 GGCATCTGTACCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556  
 571 GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 630  
 557 GGCATCTGTACCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 616  
 631 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 690  
 617 AGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676  
 691 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750  
 677 CTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736  
 751 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 807  
 737 TTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 796  
 808 ACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 867

DB 797 ATCTGAGATCCAGTGTGAGGAGCCGAGAAAGCCCTTCTCCACTCTCTCCAC 856  
 QY 868 GTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 927  
 DB 857 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916  
 QY 928 AGCAAGC 934  
 DB 917 CGCCATC 923

## RESULT 13

US-08-467-947A-1  
 ; Sequence 1, Application US/08467947A  
 ; Patent No. 6090575  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LI, YI  
 ; APPLICANT: CAO, LIANG  
 ; APPLICANT: NI, JIAN  
 ; APPLICANT: GENTZ, REINER  
 ; APPLICANT: BULT, CAROL J.  
 ; APPLICANT: SUTTON III, GRANGER G.  
 ; APPLICANT: ROSEN, CRAIG A.  
 ; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
 ; TITLE OF INVENTION: Coupled Receptor GPR1  
 ; NUMBER OF SEQUENCES: 30  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 ; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
 ; CITY: WASHINGTON  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: FLOPPY DISK  
 ; COMPUTER: IBM PC COMPATIBLE  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/467,947A  
 ; FILING DATE: 06-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/04079  
 ; FILING DATE: 30-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: STEFFE, ERIC K.  
 ; REGISTRATION NUMBER: 36,688  
 ; REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-371-2600  
 ; TELEFAX: 202-371-2540  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1713 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: both  
 ; TOPOLOGY: both  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 116..1003  
 ; US-08-467-947A-1

Query Match 3.9%; Score 121.4; DB 3; Length 1713;  
 Best Local Similarity 48.0%; Pred. No. 6.9e-20;  
 Matches 378; Conservative 0; Mismatches 406; Indels 3; Gaps 1;

151 AATGAATCCAGTGTACATCTTCAATGAGGCTCCCTGTTTGAAGAGGCTCAG 210  
 137 AATCAGACAAATGTGACAGAGTCTCTACTGAGGATTTCTCCGAGCCCAAGATTGAG 196  
 211 TTCTGTTGGCCCTTCCATGTCGCTCCCTACCTTATGCTGCTGAGTACATGACA 270